08-201: INTRODUCTION TO SOCIAL NETWORK ANALYSIS

School of Computer Science, Institute for Software Research, Carnegie Mellon University
SPRING 2010

Instructors: Kathleen M. Carley (kathleen.carley@cs.cmu.edu),
Jana Diesner (jdiesner@andrew.cmu.edu)
Terrill L. Frantz (terrill@cs.cmu.edu)

Class Time: Monday & Wednesdays, 1:30-2:20 pm

Class Location: Wean Hall 5312

Course Units: 9

Office Hours: Jana Diesner: by appointment
Terrill Frantz: by appointment

Course Description

Who is key in a group? How fast can a message spread on Facebook? Are you really six degrees away from a random stranger? Learn how to answers these questions in 08201.

Social Network Analysis (SNA) has become a widely applied method in research and business for inquiring the web of relationships on the individual, organizational and societal level. With ready access to computing power, the popularity of social networking websites such as Facebook, and automated data collection techniques the demand for solid expertise in SNA has recently exploded. In this course, students learn how to conduct SNA projects and how to approach SNA with theoretic, methodological, and computational rigor.

This interdisciplinary, undergraduate-level course introduces students to the basic concepts and analysis techniques in SNA. Students learn how to identify key individuals and groups in social systems, to detect and generate fundamental network structures, and to model growth and diffusion processes in networks. Students will be trained in interpreting the meaning of the aforementioned phenomena and suggesting potential courses of action to reinforce or change the observed trends. After this course, students will be able to design and execute network analysis projects including collecting data and considering ethical and legal implications, to perform systematic and informed analyses of network data for personal, commercial and scholarly use, and to critically review SNA projects conducted by others.

Learning Objectives

The main learning objective with this course is to enable students to put Social Network Analysis projects into action in a planned, informed and efficient manner. This overarching goal involves the following subtasks:

- Formalize different types of entities and relationships as nodes and edges and represent this information as relational data.
- Plan and execute network analytical computations.
- Use advanced network analysis software to generate visualizations and perform empirical investigations of network data.
- Interpret and synthesize the meaning of the results with respect to a question, goal, or task.
- Collect network data in different ways and from different sources while adhering to legal standards and ethics standards.

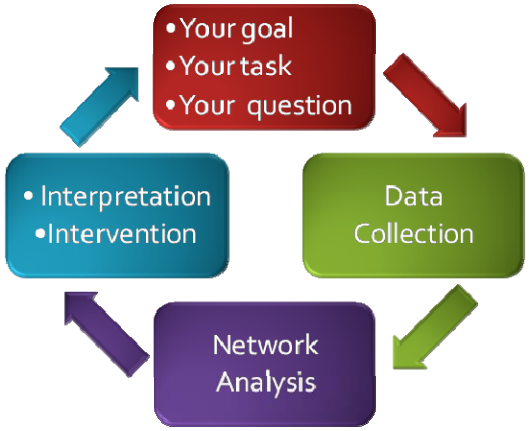
Prerequisites

This is an interdisciplinary course designed to benefit from a broad representation of students from different colleges and programs. No specific technical or numerical background is required, but students are expected to be willing to hone their computational skills. See the instructor if you have any concerns about your preparedness for this course.

Course Organization

The social network analysis process involves four basic steps as shown in the graph on the right:

1. Define a goal, question or task.
2. Collect data.
3. Analyze the data.
4. Interpret the results in order to complete a goal, answer a question, or solve a task.



In the first half of the course students will acquire the knowledge and skills needed in order to handle steps 1., 3., and 4.. In this part of the course, students learn how to investigate networks from the general to the specific, i.e. from the graph level over groups and dyads to individual nodes. For each of these levels, we will examine the observed structure by using different methods and we will interpret the meaning of the observations. Each of these levels will have a homework assignment associated with it that will be given towards the end of the section and will be due a week later.

The second half of the course serves two purposes:
First, we delve into the area of network data collection. Students will be trained in different ways of acquiring network data, including surveys, text mining, and simulations. They will also learn about the legal and ethical constraints associated with various data sources and collection techniques. This part of the course involves two homeworks.
Second, the students will put the knowledge that they acquired in part one of the course into action by planning and executing a small-scale network analysis project. The project is associated with three home work deliverables, including an in-class presentation of each team's study. This final presentation is the substitute for a final exam.

The class meeting time will be centered on lecture, but will also include a substantial amount of class discussion at times.

Materials

Software: The AutoMap and ORA software will be used through the semester. Both tools are freely available from www.casos.cs.cmu.edu Note: these software products are windows-only. They will be installed in the clusters.

Readings:

- Textbooks (required)
  - Scott, J. (2007). Social network analysis: A handbook (2nd Ed.).Newbury Park, CA: Sage.
  - Knoke (2008). Social Network Analysis,(2nd Ed).Sage.
- Other readings (required and optional) will be provided.

Evaluation and grading policy

There are six regular homework assignments. The lowest grade of the submitted regular homework assignments will be discarded.
There is one in-class mid-term, but no final exam.
Finally, there is one small-scale research project that students conduct under the guidance of the instructors. Deliverables for the project includes three homeworks.
Regular class-attendance is not graded, but strongly encouraged in order to benefit from this course.

Deliverables	**Final-grade weighting %**
6 regular homeworks	50 (10% each, lowest grad disregarded)
3 project homeworks	30 (10% each)
1 midterm	20

Course policies and expectations

Please email deliverables to the instructors on the due day prior to class. Alternatively, you can bring your submission to class and hand it in before class starts.
You are allowed one unexplained late homework in this course with a maximum submission delay of 48 hours (this delay cannot be split up among multiple home works). There is no penalty for this one late homework. For any other late homework there is a penalty of 50% grade reduction per late day.
Plagiarism and cheating are not tolerated in this course. Plagiarism means using words, ideas, or arguments from other people or sources without citation. To prevent plagiarism, cite all sources consulted to any extent (including material from the internet). Four or more words used in sequence must be set off in quotation marks, with the source identified. Cheating means copying answers from other people or sources, or providing someone with such information.
Any form of cheating will immediately earn you a failing grade for the entire course. By remaining enrolled, you consent to this policy. We will seek the harshest penalties under CMU's policy on "Standards for Academic and Creative Life" and "Cheating and Plagiarism" in the Student Guidebook (aka The Word, online at http://synergy.as.cmu.edu.edu/Student_Affairs/handbook/).

Course calendar

Date	Topic	Details	Deliverables
11-Jan	Basics	Class logistics Overview on Network Analysis	
13-Jan	Basics	The Network Analysis Process and Methodology	
18-Jan	Martin Luther King day		
20-Jan	Basics	Network Visualization	
25-Jan	Basics	When images do not suffice: Network analytical measures	HW 1 out
27-Jan	Networks: Structures, Models, Processes	Models and Simulation of Network Evolution	
1-Feb	Networks: Structures, Models, Processes	Models and Simulation of Network Evolution	
3-Feb	Networks: Structures, Models, Processes	Models and Simulation of Diffusion in Networks	HW 1 due, HW 2 out

8-Feb	Groups	Subgroups and Cliques	
10-Feb	Groups	Clustering	HW 2 due
15-Feb	Groups	Block models	HW 3 out
17-Feb	Dyads and Individuals	Ego networks	
22-Feb	Dyads and Individuals	Reciprocity	HW 3 due, HW 4 out
24-Feb	Dyads and Individuals	Social capital, structural holes, equivalence	
1-Mar	review day	bring all your questions	HW 4 due
3-Mar	in class midterm		
8-Mar	spring break, Mid-Semester Grades due		
10-Mar	spring break		
15-Mar	Ethics and Privacy		HW 5 (project) out, project data out
17-Mar	Data collection	Manual and ethnographic methods, automated methods	
22-Mar	Data collection	Cognitive Social Structures	HW 6 out
24-Mar	Project workshop 1	We as a class discuss each team's project question and provide feedback	HW 5 (project): due, HW 7 (project): out
29-Mar	Missing data: criminal networks		
31-Mar	Networks and Language	Introduction: Integration of text and network analysis	HW 6 due
5-Apr	Networks and Language	Types of networks extracted from texts across disciplines	
7-Apr	Networks and Language	Natural Language Processing and (Computational) Linguistics for Information and Relation Extraction	HW 8 out
12-Apr	Project workshop 2	Work on your project, discuss your project update and any problems with the class and the instructors	HW 7 (project): due
14-Apr	Online Communities		
19-Apr	Network over time and Simulations	Introduction: Multi-agent models for representing networks	HW 8 due
21-Apr	Network over time and Simulations		
26-Apr	review day	bring all your questions	
28-Apr	Project presentation	All project teams present a poster at an in-class poster session	Project: Poster presentation

Advice

Beyond submitting the deliverables, students will benefit greatly from the course if they participate in class discussions and discuss the topics with other students outside of class. This is a fun topic with an incredible amount of real-life application both personally and professionally no matter what life-course one takes after the semester. Social Network Analysis is still a relatively new field, so many ideas are yet unexplored. We encourage the attendees to approach this course as one that desires hard work, but to also bring an attitude of having fun. The instructors will do all they are capable of to make this an intellectually rewarding course with a good dose of fun!