CCI 688-02, MATH 410-01: INTERNET ALGORITHMS & ECON

Course Syllabus

Spring 2014 Anke van Zuylen

Basic information

- Class time & location: MWF 11:00-11:50 a.m., Jones Hall 307
- Instructor:
  Anke van Zuylen
  Jones Hall 125
  Phone: (757) 221-2036 (Office)
  Email: anke@wm.edu
  Office Hours: TR 1:30-3pm

Prerequisites:

Linear algebra, probability, computer programming (in a language of your choice), background in algorithms and datastructures is helpful but not required.

At the start of the semester, a probability review session (and potentially a linear algebra and algorithm analysis review session) will be held. Attendance at these sessions is highly recommended, even if you feel you are well prepared for this course.

Course Description:

This course covers mathematical and computational problems relating to the internet and web. We study the structure of social networks, and we explore ways to exploit these networks, for example to extract information or to advertise a product. We study game theory to model the behavior of the agents in the network, and we investigate how to design mechanisms that are robust to selfish behavior, such as online ad auctions and information routing. Finally, if time permits, we may investigate privacy and security issues relating to web data and transactions.

Text book / course materials :

The course does not have a required text book, but a highly recommended resource is the book *Networks, Crowds and Markets* by David Easley and Jon Kleinberg. The book is available online at http://www.cs.cornell.edu/home/kleinber/networks-book/networks-book.pdf. Our course will closely follow the course The Structure of Information Networks by Jon Kleinberg at Cornell (http://www.cs.cornell.edu/courses/cs6850/2011sp/) and Social and Information Network Analysis by Jure Leskovec at Stanford (http://www.stanford.edu/class/cs224w/).

Lecture notes will be posted to the course web site at blackboard.wm.edu. As part of this course, you will read papers from the scientific literature. The course web site will contain a set of resources to get you started.

Work for the course

The work for this course consists of:

- Lecture notes in LaTeX. Each student will be responsible for (i) typing up detailed lecture notes for two lectures, and (ii) carefully reviewing the lecture notes written by another student. A LaTeXtemplate will be provided. In the first week of class, we will determine the schedule. Lecture notes are due to the reviewing student by email (with a cc to the instructor) two days after the lecture, and the reviewing student has two days to review the notes. If the reviewer finds typos/omissions/mistakes, he or she can either ask the scribe to make changes, or he or she can directly make the changes. In either case, please make a list of the changes necessary, and send this to me along with the corrected notes.

- A small number of homeworks. The goal of the homework is to practice some of the skills needed for the project (exploring graph concepts, both mathematically and computationally, game theory concepts, and doing a computational study on a medium-sized data set).

- Paper presentation. To explore potential topics for the project, you will read at least two scientific papers on the same or closely linked topics, that are related to the course. The course web site will contain a list of papers to start from, but you are strongly encouraged to look for papers that are not on this list. In week 7 of the semester, you will give a class presentation on the topics you explored. Your presentation should go beyond a mere summary of the papers you read. For example, you could discuss any weaknesses you see in the paper, potential alternative hypotheses and how you would test them, or applications of the work to a different field. You are allowed to work in groups of up to three students.

- Project proposal. A project proposal of approximately 5 pages is due in week 10. The proposal should contain a clearly formulated research question, a brief discussion of relevant background work, and an outline of the steps you will follow to complete your project. You are allowed to work in groups of up to three students, and you are allowed and even encouraged to expand on the research you read for your paper presentation.

  Potential projects include mathematical modelling of some phenomenon, experimental evaluation of algorithms, empirical studies of a dataset, surveying a topic related to the course topics in much greater depth, etc. The project should contain at least some amount of mathematical analysis, and some experimentation on real or synthetic data.

  The Stanford CS224W course website at `http://www.stanford.edu/class/cs224w/projects.html` contains reports of their students' projects, as well as links to several great datasets.

- Project presentation. In the last week of classes, each group will present the results of their project.

- Final project. The final project write-up is due one week after the end of classes (although you may want to hand it in during the last week of classes, if you have a busy exam schedule in the next week).

  The project write-up is expected to be approximately 10-15 pages in length. An example of a project write-up will be distributed via blackboard.

Your grade for this course will be determined as follows: lecture notes (10%), homework (25%), paper presentation (15%), project proposal (15%), final project (35%).

At the end of the semester, I will ask each student for feedback on the contribution of their team mates on the project. This input will be used at the instructor's discretion to differentiate between students in a project group.

Class attendance and participation is a critical part of the course, since most learning will occur in class, rather than through homeworks and exams. Although class participation is not formally assigned a percentage of your total credit, you are expected to attend class and participate in class discussions.

Topics (subject to change):

- Introduction, basic graph properties, the web graph
- Small world properties in networks, decentralized search in small-world and P2P networks
- User evaluations in social media, networks with signed edges
- Cascading behavior, viral marketing, outbreak detection
- Power laws, preferential attachment models
- Link analysis for web search, Page Rank and HITS
- Spectral analysis
- Community detection and clustering
- Modeling network traffic with game theory, Braess's paradox
- Auctions, sponsored search markets

Possibility for Summer Support

The possibility exists to continue work on your project during the summer. In this case, you would still need to finish the course work, but you could define additional project goals for your summer research. For undergraduate mathematics (or double) majors, support is available from the EXTREEMS-QED program. The deadline for applications is March 17, 2014. The application form is available at `http://www.wm.edu/as/mathematics/undergraduate_research/EXTREEMS-QED/applications/application-2014.pdf`. Please come and talk to me by late February/early March if you're interested in pursuing summer research funding.

Academic Integrity:

Presenting someone else's work as your own is a violation of academic integrity, and could lead to failing the course. You are expected to maintain the highest standards of academic integrity, which means that it should always be clear which ideas you write down come from you, and which ideas were taken from others. Failing to do so is considered plagiarism. The following are a few examples:

- Verbatim copying one or more sentences from a text without citing the source;
- Verbatim copying one or more sentences from a text that is cited – verbatim copying can be done only if the sentences are clearly marked as a quotation;
- Paraphrasing one or more sentences from a text without citing the source.

You are encouraged to write your project report in LaTeX. You can then use bibtex to cite papers and create bibliographies. The course web site will contain links to information on how to use LaTeXand bibtex.