EECS6413 Information Networks

Winter 2017

Course Website
www.eecs.yorku.ca/~papaggel/courses/eecs6413/

Course Description
Information networks are effective representations of pairwise relationships between objects. Examples include technological networks (e.g., World Wide Web), online social networks (e.g., Facebook), biological networks (e.g., Protein-to-Protein interactions), and more. The study of information networks is an emerging discipline of immense importance that combines graph theory, probability and statistics, data mining and analysis, and computational social science. This course provides students with both theoretical knowledge and practical experience of the field by covering models and algorithms of information networks and their basic properties. In addition, analysis of information networks provides the means to explore large, complex data coming from vastly diverse sources and to inform computational problems and better decisions.

Topics
Topics include:

- basic graph theory
- network measurements
- network models
- community detection
- graph partitioning
- link analysis & link prediction
- information cascades & epidemics
- network ties
- recommendation systems
- mining graphs
- connections to problems in the social sciences and economics

Instructor
Manos Papagelis
Email: papaggel@eecs.yorku.ca, papaggel@gmail.com
Website: http://www.eecs.yorku.ca/ ~papaggel

Class Hours
Lectures: *Tue* and *Thu* 10:00am-11:30am at *SC 219 (Stong College)*
Office Hours: *Tue* and *Thu* 11:30am-12:30pm at *LAS 3050 (Lassonde building)*

Class Attendance
Attendance of lectures is expected but not required.

Prerequisite Courses
The course prerequisites for this course are:

- EECS-3421: Introduction to Database Systems
- EECS-3101: Design and Analysis of Algorithms
- MATH-2030: Elementary Probability
- General prerequisites

If you don't satisfy these, you need to talk with the instructor in the first week of classes to see whether you may remain in the course.

Textbooks
There is no single text for this course. The course will rely mainly on the following suggested textbooks:

- BOOK TITLE: Networks, Crowds, and Markets: Reasoning About a Highly Connected World
  AUTHORS: David Easley, Jon Kleinberg
  DATE PUBLISHED: July 2010
  FORMAT: Hardback
  PUBLISHER: Cambridge University Press
  ISBN: 9780521195331
  Note: The aforementioned book is freely available online.

- BOOK TITLE: Social Media Mining: An Introduction
  AUTHORS: Reza Zafarani, Mohammad Ali Abbasi, Huan Liu
  DATE PUBLISHED: April 2014
  FORMAT: Hardback
  PUBLISHER: Cambridge University Press
  ISBN: 9781107018853
  Note: The aforementioned book is freely available online.

- BOOK TITLE: Mining of Massive Datasets, 2nd Edition
  AUTHORS: Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman
  DATE PUBLISHED: December 2014
  FORMAT: Hardback
  PUBLISHER: Cambridge University Press
  ISBN: 9781107077232
  Note: The aforementioned book is freely available online.

- BOOK TITLE: Networks: An introduction
  AUTHORS: Mark Newman
  DATE PUBLISHED: May 2010
  FORMAT: Hardback
  PUBLISHER: Oxford Scholarship Online
  ISBN: 9780199206650

- BOOK TITLE: Social and Economic Networks
  AUTHORS: Matthew O. Jackson
  DATE PUBLISHED: November 2010
  FORMAT: Hardback
  PUBLISHER: Princeton University Press
  ISBN: 9780691148205

The course will also rely on the following reference notes/articles/reviews:

- The Structure and Function of Complex Networks (by M. E. Newman), 2003
- Structure and Dynamics of Information in Networks (by David Kempe), 2011

In addition, a number of recent research papers in the area of information networks will be distributed in every iteration of the class.

Communication

The main communication tools for the class will be the *course website* and *Piazza*.

- *Course Website*: Most class materials are available on the course web site; be sure to check regularly. The page has also a link to a discussion board. We are using Piazza.
- *Piazza*: Instead of a discussion board, we are using Piazza, a free Q&A platform. Piazza can get you fast, accurate response to your questions – but it only works if everyone participates! We will also use Piazza to post announcements and updates, so both the website and Piazza is required reading. See below for Piazza signup and class links:

  Signup link: www.piazza.com/yorku.ca/winter2017/eecs6413

  Class link: www.piazza.com/yorku.ca/winter2017/eecs6413/home

  Note: You will need to sign up with your school email, ending in *yorku.ca*. If you do not have a school email address, please contact your instructor and request to be enrolled with your personal email.
- *Email*: Please use email for personal issues and the discussion board to ask general course-related questions. Include "eecs6413" in all email subject lines to ensure your message is correctly filtered and filed. An informative subject line like "eecs6413: Question related to X" really helps. I try to respond to email frequently. However, due to volume, it may take longer, especially on weekends and near due dates.

Grading Policy

Work	Weight	Comment
2 Assignments	30%	15% each
Project	40%	Large project consisting of: proposal, milestone report, final presentation, and final report
Final Exam	30%	You must get >=40% to pass the course

Final Examination

A written final exam will be given between 7-24 Apr (to be determined during the term).

Working with a Partner

You have the option of partnering with one other (currently enrolled) students for your assignments, and we encourage you to do so. The ability to work effectively in a team will be very important in your career, and that involves many skills beyond the purely technical aspect of creating working code. You may choose your own partner, and it need not be the same person for each assignment. If you do have a partner for an assignment, submit only a single copy of your work. Jointly submitted assignments will be graded in the usual way and both partners will

receive the same mark. Working with a partner has the potential to lighten your workload or to increase it, depending on how well you work together. Be aware that simply dividing the work and assembling your separate pieces at the end is a poor strategy for completing successful assignments. And of course, you are responsible for learning the course material underlying all parts of the assignments. You will have the most success if you truly work together.

Assignment Policies
You must make sure that all your assignments are running and are sufficiently documented. Code that doesn't compile, fails to run or lacks documentation, will be marked as not working.

Late Work Policy
The late policy is strict. All assignments will be submitted electronically. Late assignments will be handled based on a system of "*grace days*", as follows: Each student begins the term with 3 *grace days*. One grace day is 24 hours. If an assignment is due at 10:00 p.m. on a Friday then an assignment handed in by 10:00 p.m. on Saturday uses one grace day. The grace days are intended for use in emergencies (e.g., system failure or illness). Do not use all of them to buy an extension because of a busy week or you will be out of luck in a true emergency. Assignments submitted after the due date when all grace days have been used will receive a grade of 0.

If you are at risk of missing a deadline due to a busy week, rather than using your grace days you should hand in a working (and tested) version of a simpler program. In the event of an illness or other catastrophe, get proper documentation (e.g., medical certificate), and contact me (by email or in person) as soon as possible. Do not wait until the due date has passed. It is always easier to make alternate arrangements before the due date or test day.

Assignments are submitted electronically and will often be tested using an automated testing program; you must follow the submission instructions exactly. If you do not, you will most likely lose substantial marks on the assignment. If you find you have submitted the wrong file or omitted a file, please notify your instructor as soon as possible.

Remarking
If you feel an error was made in marking an assignment or test please submit a remark request. Requests for remarking must be submitted using a university remarking request form explaining what your concern is **no later than a week after** the assignment (or test) has been returned back.

Academic Offenses
All of the work you submit must be done by you and your work must not be submitted by someone else. Plagiarism is academic fraud and is taken very seriously. The department uses software that compares programs for evidence of similar code. Please read the Rules and Regulations from the York University's Academic Integrity and the York University's Senate Policy on Academic Honesty documents.

Accessibility Needs
York University is committed to accessibility. If you require accommodations for a disability, or have any accessibility concerns about the course, the classroom or course materials, please contact York University's Counselling & Disability Services.

Tentative Schedule
A tentative schedule of topics to be covered appears below. This is subject and likely to change.

Lecture Schedule

Week	Week of	Topic
1	2 Jan	**Lecture 1: Introduction** Introduction to main problems about networks. Basic mathematical concepts.
2	9 Jan	**Lecture 2: Network Measurements** Degree distributions. Measuring power-laws. Clustering Coefficient, Effective Diameter, Bow-tie structure, Homophily.
3	16 Jan	**Lecture 3: Network Models** Erdos-Renyi graphs. Configuration Model. Preferential Attachment. Small-world models. Forrest-Fire model. Kronecker graphs.
4	23 Jan	**Lecture 4: Community Detection** Communities in Social Networks, Clustering, Betweenness, Modularity.
5	30 Jan	**Lecture 5: Graph Partitioning, Densest Subgraph** Graph Partitioning, Spectral Clustering. The Densest Subgraph problem.
6	6 Feb	**Lecture 6: Link Analysis** Web search, PageRank, HITS. Random walks on graphs. Absorbing Random Walks. Opinion diffusion.
7	13 Feb	**Lecture 7: Link Prediction** Link prediction and link recommendations.
8	20 Feb	**No Classes** – Reading week and family day
9	27 Feb	**Lecture 8: Information Cascades** Information Cascades, Epidemics, Influence Maximization, Game theoretic information cascade. Models for epidemic spread.
10	6 Mar	**Lecture 9: Network Ties** Strong and Weak ties. Strong Triadic Closure. Networks with Positive and Negative ties. Structural Balance.
11	13 Mar	**Lecture 10: Topics in Mining Information Networks** (e.g., Team Formation in Social Networks, Social Recommendations, etc.)
12	20 Mar	**Lecture 11: Topics in Mining Social Content** (e.g., Using content from online social networks and media to predict stock changes, track earthquakes, and understand news cycles.)
13	27 Mar	**Lecture 12: Student Presentations** Project presentations
14	3 Apr	**Lecture 13: Student Presentations and Wrap Up** Project presentations, wrap up, reviews

Tentative Assignment Schedule (30%)

Assignment	Weight	Delivery Date	Due Date	Topic
1st	15%	TBD	TBD	TBD
2nd	15%	TBD	TBD	TBD