ECE 442 - Network Science Analytics

Syllabus - Spring 2023

Time: Mondays and Wednesdays, 3:25 pm - 4:40 pm.
Place: Computer Studies Building 601.
Class website: http://www.hajim.rochester.edu/ece/sites/gmateos/ECE442.html

Instructor: Gonzalo Mateos (gmateosb@ece.rochester.edu).
Office hours: Tuesdays 2 pm in CSB 726.
Teaching assistant: Narges Mohammadi (nmohamm4@ur.rochester.edu).
TA office hours: Thursdays 2 pm in CSB 633.

Bibliography:

- Eric D. Kolaczyk, *"Statistical Analysis of Network Data: Methods and Models,"* Springer.

Available online from the University of Rochester library.

Additional reading:

- M. E. J. Newman, *"Networks: An Introduction,"* Oxford University Press.
- D. Easley and J. Kleinberg, *"Networks, Crowds, and Markets: Reasoning About a Highly Connected World,"* Cambridge University Press.
- J. Leskovec, A. Rajaraman and J. D. Ullman, *"Mining of Massive Datasets,"* Cambridge University Press.
- Research papers provided by the instructor.

The books will be on reserve for the class in Carlson Library.

Prerequisites: Useful to have good background in probability theory and linear algebra, as well as some exposure to graphs. For homework assignments we will use Matlab/Python and possibly NetworkX as well as PyTorch Geometric.

Credit distribution: Homework assignments (3-5, 30%), and a research project in groups of two students involving three deliverables (proposal 10%, progress report 10%, final report and in-class presentation 50%). Detailed information on the project will be posted in the class website.

Academic dishonesty: Academic dishonesty will be dealt with according to the University of Rochester's Academic Honesty Policy.

Class description: Network Science Analytics (ECE 442) is a graduate class about networks. The science of networks is an emerging discipline of great importance that combines graph theory, probability and statistics, and facets of engineering and the social sciences. Topics in this course will help answer intriguing questions such as: Where does "six degrees of separation" come from? How can we make sense of large graphs, ranging from social networks to the smart power grid? What are the underpinnings of Google's search engine and webpage ranking? What are good models for predicting popularity in Twitter? How can we estimate the size of the Internet?

Class objectives: This course will provide students with the mathematical tools and computational training to understand large-scale networks in the current era of Big Data. It will introduce basic network models and structural descriptors, network dynamics and prediction of processes evolving on graphs, modern algorithms for topology inference, community and anomaly detection, as well as fundamentals of social network analysis. All concepts and theories will be illustrated with numerous applications and hands-on case studies from technological, social, biological, and information networks.

Topic outline: Topics covered in ECE 442 will tentatively include

1) Introduction and overview. Why study networks?
2) Preliminaries: background on graphs, probability, and optimization theory.
3) Descriptive analysis of networks: centrality, cohesion, and communities.
4) Fundamentals on social network analysis.
5) Mapping networks, visualization of large graphs.
6) Sampling and estimation in network graphs
7) Network models: random, small-world, preferential attachment.
8) Network topology inference: link prediction, tomographic inference.
9) Modeling and prediction for processes evolving over network graphs.
10) Analysis of network flow data.

For a detailed description of the course contents including a lecture-by-lecture schedule, visit the class website at http://www.hajim.rochester.edu/ece/sites/gmateos/ECE442.html.

Tentative class schedule - Spring 2023

Date	Description	HW/Project
Wed. 01/11	Traveling to DARPA QuICC kick-off meeting - No class	
Fri. 01/13	Introductions, class organization, networks, context, examples	
Mon. 01/16	Martin Luther King Jr. Day - No class	
Wed. 01/18	Traveling to McGill Bellairs Workshop - No class	
Mon. 01/23	Graphs, digraphs, degrees, movement, strong and weak connectivity	
Wed. 01/25	Families, algebraic graph theory, data structures and algorithms	
Fri. 01/27	Inference, models, point and set estimates, hypothesis testing	
Mon. 01/30	Tutorials on inference about a mean and linear regression	
Wed. 02/01	Graph visualization, stages of network mapping, mapping Science	
Mon. 02/06	Large graph visualization, k-core decomposition, Internet mapping	
Wed. 02/08	Degree distributions, Erdos-Renyi random graphs and power laws	HW1 due
Mon. 02/13	Visualizing and fitting power laws, preferential attachment	
Wed. 02/15	Closeness, betweeness and eigenvector centrality measures	
Mon. 02/20	Web search, hubs and authorities, Markov chains review	
Wed. 02/22	PageRank, fluid and graph random walk models, distributed algorithms	
Mon. 02/27	Cohesive subgroups, clustering, connectivity, assortativity mixing	
Wed. 03/01	Strength of weak ties, community structure in networks	Proposal
Mon. 03/06	Spring Break - No class	
Wed. 03/08	Spring Break - No class	
Mon. 03/13	Girvan-Newmann method, hierarchical clustering, modularity	
Wed. 03/15	Modularity optimization, graph cuts, spectral graph partitioning	
Mon. 03/20	Sampling, Horvitz-Thompson estimation, graph sampling designs	
Wed. 03/22	Network estimation of totals, groups size, degree distributions	HW2 due
Mon. 03/27	Random graph models, model-based estimation, significance, motifs	
Wed. 03/29	Small-world, preferential attachment and copying models	
Mon 04/03	Exponential random graph models, construction and estimation	Prog. Report
Wed. 04/05	Topology inference, link prediction, scoring and classification	
Mon. 04/10	Inference of association networks, tomographic inference	
Wed. 04/12	Nearest-neighbor prediction of processes, Markov random fields	
Mon. 04/17	Graph kernel-regression, kernel design, protein function prediction	
Wed. 04/19	Diseases and the networks that transmit them, epidemic modeling	HW3 due
Mon. 04/24	Network flow data, routing and traffic matrices, gravity models	
Wed. 04/26	Traffic matrix estimation, network flow costs, network kriging	
TBD	In-class student project presentations	Presentation