CptS 591: Elements of Network Science

Semester Project & Mid-Term Exam

Assefaw Gebremedhin
School of Electrical Engineering and Computer Science
assefaw.gebremedhin@wsu.edu
http://www.eecs.wsu.edu/~assefaw

Topics

Covered:

- Introduction
- Graph theory refresher
- Basic network properties (2)
- Intro to igraph + network vis (2)
- Random graphs (2)
- Spectral analysis (2)
- Centrality
- Link analysis: PageRank
- Link analysis: Hubs and Authorities
- Signed networks

To come:

- Community identification (clustering)
- Graph similarity with (un)known node correspondence
- Graph embedding
- Cascading behaviors
- Influence maximization
- Epidemic models
- Time varying networks

Project

- Constitutes 50% of final grade
- Team of 2 or 3
- Topic and scope: largely up to you as long as the project
  - Clearly falls within the realm of the course
  - Is likely to be doable within the remainder of the semester
  - Is reasonably novel and interesting (e.g. could form a basis for further research)

Project: four basic types

- *Analysis* of an interesting dataset using existing implementations of algorithms, models and measures.
- *Implementation* of a new algorithm, model or measure and evaluating its performance on a collection of datasets.
- A *theoretical* project that considers an algorithm, a model or a measure and derives rigorous results about it.
- A *critical survey* of a specific topic in an area of the course, going in some depth and offering a fresh perspective.

A project may also (and is encouraged to) be a combination of some of the above.

Project Parts

- Reaction Paper
  - Due March 29
- Project Proposal
  - Due April 5
- Final Report
  - Due May 3
- Presentation (12-min presentation + 3-min Q&A)
  - Scheduled for April 22, 27, 29

Project parts: Reaction Paper

- You get to pick two closely related papers to read
- Produce a reaction paper (2 to 3 pages) that contains:
  - Summary:
    - What is the main technical content of the papers?
    - How do the papers relate to the topics of the course?
    - What is the connection between the papers?
  - Critique:
    - What are the strengths and weaknesses of the papers?
    - Were there any unrealistic assumptions made?
  - Further work:
    - What are some promising further research questions along the directions of the papers?
    - How can the works be extended? Do you see an idea for a better model? A better algorithm?
    - Another problem or data the methods can be applied to?
- One of the goals is to help you generate a project idea for your proposal. (If you already have an idea, then you may choose the papers you read so that they have some relationship with your idea.)

Project parts: Proposal

- Ideally an outgrowth of the further work you identified in the Reaction Paper, but it could also be detached from it.
- **Can also be chosen from a list of ideas I will provide.**
- Roughly 2 to 3 pages.
- It should accomplish the following
  - Describe what you intend to do
  - Describe the methods you plan to use
  - Describe the data you will use and discuss how you plan to obtain it
  - Discuss relevant background work
  - Discuss your tentative plan

Project parts: Final Report

- A report of about 8 to 15 pages
- Content should roughly map to
  - *Introduction/Motivation/Problem Definition*
    - (where you state what you are trying to solve/achieve and why it matters)
  - *Model/Algorithm/Method*
    - (where you give a detailed description of your work)
  - *Results and findings*
    - (where you interpret the results you obtain, discuss implications, make observations and draw conclusions)
  - *Related work*
    - (where you cite and briefly summarize other work related to yours.)
- Later in the semester, I will provide a more detailed guideline on how to write an effective final report.

Datasets (may also inspire project ideas)

- Stanford Large Network Dataset Collection
- Co-authorship and citation networks
- Internet topology
- Stack overflow
- Yelp data
- Peer to peer money lending dataset
- YouTube dataset
- Amazon product copurchasing data
- Wikipedia
  - Page-to-page link data Dbpedia Edit history
- Movie ratings
- Who trusts whom data at Trustlet
- Interesting websites (for analysis):
  - Delicious Digg reddit StumbleUpon The Hype Machine Twitter jyte.com prosper.com lendingclub.com mturk.com

- http://snap.stanford.edu/data/
- http://www-personal.umich.edu/~mejn/netdata/

Software tools

- igraph (main tool used in the course)

Other (potentially relevant) tools:

- NetworKit (https://networkit.iti.kit.edu)

- SNAP (C++, Windows, Linux)
- Networkx
- Gephi (visualization software)

- GUESS (exploratory data analysis and visualization)
- Infovis cyberinfrastructure (Linux, Windows, MacOSX)

Mid-term

May cover highlighted (blue) topics:

- **Introduction**
- **Graph theory refresher**
- **Basic network properties**
- **Intro to igraph**
- **Random graphs**
- **Spectral analysis (part II)**
- **Centrality**
- **Link analysis: PageRank**
- **Link analysis: Hubs and Authorities**
- **Signed networks**

Other matters:

- **Recall that mid-term accounts for 18% of final grade**
- **It is a written exam**
- **Take-home with online (Canvas) submission. Due within 24-hours after release**
- **The exam will consist of "Problem" type questions, requiring some ingenuity, as opposed to "Exercise" type questions that would review definitions and concepts**
- **Potential exam dates (open for vote):**
  - **March 18**
  - **March 23**
  - **March 25**

Schedule we had at semester start and updates

Week	Topics	Assignments
01 (Jan 19/21)	Introduction, Graph theory refresher	Survey out
02 (Jan 26/28)	Network properties	Survey due, Assignment 1 out
03 (Feb 02/04)	Intro to igraph	
04 (Feb 09/11)	Random graphs	Assignment 1 due
05 (Feb 16/18)	Spectral graph theory	Assignment 2 out
06 (Feb 23/25)	Centrality	NO CLASS 2/25
07 (Mar 02/04)	PageRank, Hubs & Authorities	Assignment 2 due
08 (Mar 09/11)	Community detection, project discussion	
09 (Mar 16/18)	Graph similarity, signed networks	Mid-Term
10 (Mar 23/25)	Graph embeddings	Reaction paper due
11 (Mar 30/Apr 01)	Cascading behaviors	Project proposal due
12 (Apr 06/08)	Influence maximization, Epidemic models	Assignment 3 out
13 (Apr 13/15)	Temporal networks	Assignment 3 due; NO CLASS 4/13
14 (Apr 20/22)	Wrap-up, project presentations	
15 (Apr 27/29)	Project presentations	
16 (May 04/06)	Finals week	Final project report due

Updates: likely no Assignment 3;
minor updates in order of topics, including in weeks 8 and 9