# CptS 591: Elements of Network Science

***Review &Wrap Up***

# Goals (what we set out to do…)

Students will be introduced to select

- mathematical and computational methods used to *analyze* networks
- models used to understand and *predict* behavior of networked systems
- theories used to reason about network *dynamics*

And students will apply what they learn by completing a semester project and a set of assignments

# Topics we covered

1. Graph theory (one lecture)
   Nodes and edges. Paths. Cycles. Connectivity. Components. Distance. BFS.
   Material: lecture slides + Chap 2 of Easley-Kleinberg.
2. Basic network properties (two lectures)
   Degree distribution. Path lengths and distributions. Clustering coefficient.
   Material: lecture notes + slides.
3. Random graphs (two lectures)
   Random graphs as a concept. Random variables and expectation.
   Graph invariants in random graphs. Phase transition.
   Random graphs vs real-world networks.
   Material: lecture notes + slides.
4. Spectral analysis (two lectures)
   Basic linear algebra review.
   Spectrum of three matrices associated with a graph:
   the adjacency matrix, the Laplacian and the normalized Laplacian.
   The second eigenvalue of the Laplacian and its significance.
   Isoperimetry. Spectra of subgraphs and supergraphs.
   Material: lecture notes.

# Topics we covered

5. Centrality (two lectures)
   Elementary common denominator formalization.
   Centrality around distances and neighbors. Centrality around shortest paths.
   Feedback centrality (Katz Index).
   Material: lecture slides + lecture notes.

6. PageRank (one lecture)
   Random surfer derivation of PageRank. Markov chains.
   Solving the PageRank vector system.
   Material: lecture notes.

7. Hubs and Authorities (one lecture)
   Hub score, Authority score, and the HITS algorithm.
   Material: Chap 14 of Easley-Kleinberg.

8. Signed networks (one lecture)
   Structural balance.
   Trust and distrust.
   Weakly balanced networks.
   Balance in general networks.
   Material: lecture slides + Chap 5 of Easley-Kleinberg

# Topics we covered

9. Community identification (two lectures)

Overview of community detection paradigms and algorithms

Minimum-cut. Edge Betweenness. Modularity maximization (Louvian).

Label propagation.

Overview of methods available in igraph.

Direction-optimizing label propagation (DOLPA) algorithms and applications

Material: lecture slides

10. Similarity between nodes in a graph (one lecture)

Structural equivalence – cosine similarity. Pearson coefficients.

Regular equivalence (Katz similarity).

Material: lecture slides + sec 7.2 of Newman.

11. Similarity between graphs and applications (one lecture)

Graph embeddings.

Factorization-based methods. Random walk approaches. Encoder-decoder architectures. Graph convolutional based approaches.

Graph similarity with known node correspondence.

Similarity with unknown node correspondence.

Material: lecture slides + ICDM 2014 tutorial

# Topics we covered

12. Cascading behaviors (one lecture)
    Diffusion in networks. Cascades and clusters.
    Thresholds and the role of weak ties. Heterogenous thresholds.
    Collective action and pluralistic ignorance.
    Material: lecture slides + Chap 19 of Easley-Kleinberg
13. Influence maximization (one lecture)
    Viral marketing.
    Independent Cascade Model. Linear Threshold Model.
    Maximizing spread of influence under ICM and LTM.
    Hill-climbing algorithm. Nemhauser-Wolsey-Fisher theorem. Submodularity.
    Material: lecture slides + Kempe-Kleinberg-Tardos KDD03 paper.
14. Epidemics (one lecture)
    Branching processes.
    The SIR model. The SIS model.
    Synchronization. Transient contacts and concurrency. Genetic inheritance.
    Analysis of branching and coalescent processes.
    Material: lecture slides + Chap 21 of Easley-Kleinberg.

# Other lectures incorporated in the semester

- igraph (network analysis software)
  - Tutorials (two lectures)
- Semester project setup discussion
  - (one lecture)
- Future of the US power grid
  - Webinar

# Pointers to a couple things we didn't quite cover…

- Power-Laws
  - See the tutorial by Lada Adamic titled Zipf, Power-laws and Pareto:

  http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html
- Chris Anderson's 2004 Wired-article titled The Long Tail

  https://www.wired.com/2004/10/tail/
- Small-world phenomenon
  - Milgram's experiment (six degrees of separation)
  - Watts-Strogatz model (Nature, 1998)
  - Kleinberg's decentralized search model
  - Material: Chapter 20 of Easley-Klienberg
- Community detection in networks: A user guide (S. Fortunato and D. Hric, Physics Reports 2016). https://www.sciencedirect.com/science/article/pii/S0370157316302964#s000115
- Temporal networks
- For a good overview of temporal network mining models and algorithms, visit this KDD 2019 tutorial:

  https://dl.acm.org/doi/10.1145/3292500.3332295

# Assignments, Home works and Mid-term

- Assignment 1
  - Basic network analysis (using igraph)

    * Calculating basic network properties.
    * Plotting degree distributions.
    * Plotting pathlength distributions.
- Assignment 2
  - Spectral analysis and centrality (using igraph )

    * Calculating various centrality measures
      (degree, PageRank, Authority score, Hub score, betwenness, closeness, eccentricity).
    * Generating synthetic random graphs (Erdos-Renyi, Barabasi-Albert).
    * Calculating eigenvalues and eigenvectors of the Laplacian. * Plotting eigenvectors.
- Homework 1
  - Graph theory (paper-and-pencil problems; in-class submission)
- Homework 2
  - Quick reflections on the future of the US power grid webinar
- Mid Term (take-home)
  - Network properties, Affiliation networks , Hubs-And-Authorities Algorithm, Spectral Graph Theory

# Semester Project

- Reaction Paper
  - Aim: read a couple of research papers on a chosen course topic(s) and provide a 2-3 page systematic critique
- Project
  - Proposal
  - Presentation
    - Scheduled for 4/27, 4/29
  - Final Report
    - In: 5/3 by 11:59pm

# Guidelines (details posted on Canvas,

please read the guidelines before starting to prepare your report/presentation)

- Final Report
  - Length: 10 – 20 pages (excluding appendix)
  - Content:
    - Abstract
    - Introduction
    - Problem Definition
    - Models/Algorithms/Measures
    - Implementation/Analysis
    - Results and Discussion
    - Related Work
    - Conclusion
    - Bibliography
    - Appendix (code listing + pointer to electronic copy of code)
- Presentation
  - Length: 8 min talk + 2 min discussion
  - Format: flexible
  - Remember to provide background and context (for your classmates)

# Project presentations schedule

**Tuesday, April 27**

1. **Team 1**
   Probabilistic synthesis of temporal graphs
2. **Team 5**
   Adversarial social interaction: case study on MMO Eve Online dataset
3. **Team 13**
   Game-theoretic analysis of information campaigns on Twitter networks
4. **Team 6**
   Community detection for product recommendation
5. **Team 11**
   Community detection for analysis of Bitcoin networks
6. **Team 8**
   Information flow analysis on Bitcoin datasets

**Thursday, April 29**

1. **Team 3**
   Analysis of hospital transfer networks
2. **Team 4**
   Semi-supervised learning using graph signals
3. **Team 2**
   Analysis of subreddits with a view on conflicts and attacks
4. **Team 7**
   Core-Periphery structure analysis of Bitcoin data
5. **Team 9**
   Graph similarity
6. **Team 10**
   Analysis of geographic networks of Covid data
7. **Team 12**
   Improving collaborative filtering recommender system using optimization

# Grading reminder…

Final Grade:

- Project: 50%
    - Reaction paper: 6%
    - Project proposal: 4%
    - **Presentation: 10%**
    - **Final report: 30%**
- Assignments: 30%
    - Assignment 1: 13%
    - Assignment 2: 13%
    - Homework 1: 2%
    - Homework 2: 2%
- Mid Term: 18%
- Class participation: 2%

# Learning outcomes we set out to achieve…
(check each item and see if you would give yourself a nod)

- Explain basic metrics and measures used to characterize networks
- Analyze a network using the various measures and a suitable network analysis software tool
- Discuss the strengths and weaknesses of random graph models
- Understand and apply key algorithms for node ranking, community identification, and network similarity
- Understand and apply models and theories used to reason about cascading behaviors, information diffusion, contagion, decentralized navigation
- Understand and explain the interdisciplinary nature of network science
- Critique research papers in the area
- Apply knowledge gained in the course to carry out a project and write a scientific report

# Thanks!

- And good luck with your projects and your presentations and final reports!