Information Networks
Review of Key Concepts
The “Age of Networks”

Technological  Social  Biological
why should we care about networks?
Why Networks? Why Now?

- **Universal language for describing complex data**
  - Networks from science, nature, and technology are more similar than one would expect

- **Shared vocabulary between fields**
  - Computer Science, Social science, Physics, Economics, Statistics, Biology

- **Data availability (computational challenges)**
  - Web/mobile, bio, health, and medical

- **Impact!**
  - Social networking, Social media, Brain, Drug design
  - We will never understand these systems unless we understand the networks behind them!
how do we reason about networks?
How do we reason about networks?

- **Empirical**: Study network data to find organizational principles
- **Mathematical models**: Probabilistic, graph theory
- **Algorithms**: Methods for analyzing graphs
What do we study in networks?

- **Structure and evolution**
  - What is the structure of a network?
  - Why and how did it become to have such structure?

- **Processes and dynamics**
  - Networks provide “skeleton” for spreading of information, behavior, diseases
What Have We Covered?

- **basic graph theory**
  - graphs, networks
  - bow-tie structure
- **network measurements**
  - degree distributions, power-laws
  - shortest paths, clustering coefficient
- **network models**
  - Erdos-Renyi model
  - small-world model
  - configuration model
  - scale-free networks
- **models of evolving graphs**
  - preferrential attachment model
  - microscopic/macroscopic evolution of networks
  - forest-fire model
- **community structure in networks**
  - Strength of weak ties, structural holes
  - community detection, Girvan-Newman algorithm
  - graph partitioning, graph cuts, conductance
  - spectral graph theory, spectral graph clustering
- **overlapping communities in networks**
  - cliques, clique percolation method
  - community-affiliation graph model
- **link analysis**
  - web search
  - hubs and authorities (HITS)
  - PageRank, topic-sensitive PageRank
- **link prediction**
  - neighborhood-based methods
  - node proximity based methods, supervised learning models, Facebook's "PYMK", Twitter's "WtF"
- **cascading behavior in networks**
  - Granovetter’s model, threshold model
  - game theoretic model
  - epidemic model on trees
  - disease spreading models (SIR, SIS, SIRS)
  - independent cascade model
  - influence maximization
  - outbreak detection
- **Advanced Topics: ML with Graphs**
  - network representation learning
  - graph neural networks
  - reasoning over knowledge graphs
  - applications of ML with graphs
Small-World Phenomena

- **Properties:**
  - Six degrees of separation
    - Networks have small diameters
  - Edges in the networks cluster
    - Large clustering coefficient

- **Models:**
  - Erdös-Renyi model
    - Baseline model for networks
  - The Small-World model
    - Small diameter and clustered edges

- **Algorithms:**
  - Decentralized search in networks
    - Kleinberg’s model and algorithm

\[ C_1 = \frac{1}{3} \]

\[ P(u \rightarrow v) \sim d(u, v)^{-\alpha} \]
Scale-Free Networks

- **Properties:**
  - **Power-law degrees**
    - Degrees are heavily skewed
  - **Network resilience**
    - Networks are resilient to random attacks

- **Models:**
  - Preferential attachment
    - Rich get richer

- **Algorithms:**
  - Hubs and Authorities
    - Recursive: \( a_i = \sum_{j \rightarrow i} h_j, \quad h_i = \sum_{i \rightarrow j} a_j \)
  - PageRank
    - Recursive formulation, Random jumps
Community Detection

- **Properties:**
  - Strength of weak ties
  - Core-periphery structure

- **Models:**
  - Community-affinity model

- **Algorithms:**
  - Spectral Clustering
  - Girvan-Newman (Betweenness centrality)
  - **Modularity:** \#edges within group – E[\#edges within group]
  - Clique Percolation Method
    - Overlapping communities
Network Evolution

- **Properties:**
  - Densification Power Law
    - $E(t) \propto N(t)^a$
  - Shrinking Diameter

- **Models:**
  - Microscopic Network Evolution
    - Exponential life-times, Evolving sleeping times
    - Random-Random edge attachment

- **Algorithms:**
  - Link prediction
Network Diffusion (1)

- **Properties:**
  - Node-to-node influence
  - Node threshold
  - Cascade spread

- **Models:**
  - Game theoretic model:
    - Payoffs, Competing products
  - Independent Cascade Model
    - Each node infects a neighbor with some probability

---

![Diagram of network diffusion with node influences and thresholds]
Algorithms:

- **Influence Maximization**
  - Set of $k$ nodes producing largest expected cascade size if activated
  - Submodularity
  - Greedy hill-climbing

- **Outbreak Detection**
Map of Superpowers

Properties
- Small diameter, Edge clustering
- Scale-free
- Strength of weak ties, Core-periphery
- Densification power law, Shrinking diameters
- Information virality, reproductive number

Models
- Small-world model, Erdös-Renyi model
- Preferential attachment, Copying model
- Community-affiliation Graph Model
- Microscopic model of evolving networks
- Independent cascade model, Game theoretic model, SIR

Algorithms
- Decentralized search
- PageRank, Hubs and authorities
- Community detection: Girvan-Newman, Modularity
- Link prediction, Supervised random walks
- Influence maximization, Outbreak detection, LIM
Applying Your Superpowers
Applying Your Superpowers

- Social media analytics
- Viral marketing
Applying Your Superpowers

- Predicting epidemics: Ebola
Interactions of human diseases

Drug design
What’s Next?
What’s Next?

- **Project presentation**
  - Tue, Dec 8th, in-class
    - 9 minutes + 2 min QA
    - See course website for more info

- **Project final report**
  - Sun, Dec 20th Midnight (11:59PM)
    - see course website for more info

- **Final exam**
  - Tue, Dec 22nd, 9am-12pm
    - Short answers
    - Online (EClass)
What Next? Seminars

- **EECS6xxx: Data Analytics and Visualization**
  - Winter 2021, Project course
  - Data mining, graph mining, data visualization

- **Conferences / Journals:**
  - **Conferences**
    - **KDD**: Conf. on Knowledge Discovery & Data Mining
    - **WWW**: ACM World Wide Web Conference
    - **WSDM**: ACM Web search and Data Mining
    - **ICDM**: IEEE International Conference on Data Mining
    - **ICWSM**: AAAI Int. Conf. on Web-blogs & Social Media
  - **Journals**
    - **Complex Networks**: Journal of Complex Networks
    - **TKDD**: ACM Transactions on Knowledge Discovery from Data
    - **TKDE**: IEEE Transactions on Knowledge and Data Engineering
In Closing...

You have worked a lot...

...and (hopefully) learned a lot!
thank you & happy holidays