Graph Theory and Social Networks



What is a Graph?

- Can you draw one?
- How does your graph represent:
 - the whole data set?
 - an individual data point?
 - the relation between two points?
- What do all graphs attempt to convey?

RELATIONSHIPS



Graph Basics

In its most basic form, a graph is an ordered pair G = (V, E)
V is a set of vertices (or nodes or points)
E is a set of edges (or arcs or lines), a 2-element subset of V
Can you draw this graph G = (V, E) ?
V = {1,2} E = {(1,2)}
This graph is both undirected and simple

Simple Graph and Multigraph



Directed and Undirected Graphs





Graph Properties

Order (|V|) - the number of vertices in a graph

Vertex Degree

the number of edges connected to a vertex

Centrality

which nodes are most _ important?

Graph Centrality



Given graph G = (V, E), and some vertex $v \in V$

Degree Centrality: Your connections are what matter.

 $C_D(v) = \deg(v)$

Graph Centrality



Given graph G = (V, E), and some vertex $v \in V$

Degree Centrality: Your connections are what matter.

 $C_D(v) = \deg(v)$

Closeness Centrality: Your distance from everyone else in the network is what matters.

$$C_{\ell}(v) = \frac{1}{\sum_{u \in V} d(v, u)}$$

Graph Centrality



Given graph G = (V, E), and some vertex $v \in V$

Degree Centrality: Your connections are what matter.

 $C_D(v) = \deg(v)$

Closeness Centrality: Your distance from everyone else in the network is what matters.

$$C_{\ell}(v) = \frac{1}{\sum_{u \in V} d(v, u)}$$

Betweenness Centrality: Your position as an intermediary in the network is what matters.

$$C_B(v) = \sum_{s,t \in V} \frac{\sigma_{(s,t)}(v)}{\sigma_{(s,t)}}$$



Social Network Analysis

- What is a Social Network?
 - Can you draw yours?
- Activity: Draw the social network of *your* dorm floor
 - Things to consider:
 - What constitutes an edge?
 - Is your network directed?
 - Are there any bridges? Any connected components?











Middle-High School

Are our lives structured?



Why does it matter?

- Which came first- social identity/behavior or the structures that allow them to occur?
- Social networks inscribe everyone
 - And their structure results in unintuitive findings!
 - Kevin Bacon number
 - 🎶 It's a small world after all... 🎶

Milgram's Experiment

- Waaay back in 1969, Stanley Milgram (at Columbia!) showed that people are cognizant of the correlation between network connections and demographic characteristics



Algorithmic Influence Propagation



- The "small world" structure of human relations potentially allows for rapid, widespread communication
 - Of ideas
 - Of disease
 - Of products
 - In industry, marketing divisions often turn to network analysts for recommendations on how to insure the widespread of a product
 - Ever heard of FitTea?

Algorithmic Influence Propagation

- Who should we get to promote our fitness goalz?
 - Who is the most central node?
 - Does the structure effect edge importance?





Algorithmic Influence Propagation

- Can we approach this question from an algorithmic perspective?



Top-k Degree

topK(G, k)

for each vertex $v \in V$ do $dd_v = deg(v)$ end for seedSet = findMax(dd_v, k) output seedSet. Will this give us the best result?

Who do you go to for product recommendations?

- How can we encode the "importance" of our relations?
- Is this a node, or structural, attribute?

Greedy Node Selection



Figure 5: Influence spreads of different algorithms on the collaboration graph NetHEPT under the weighted cascade model (n = 15, 233 and m = 58, 891).

Will this give us the best result?

Will this always give us the same result?

Is this algorithm sensitive to socio-demographic characteristics? -Implicitly or explicitly?

At Columbia



Vishal Mishra

Augustin Chaintreau

David Stark



Courses to Consider: COMS W4995 -- Networks and Crowds COMS W6998 -- Social Networks COMS W6998 -- Internet Economies SOCI C4701 -- Social Network Analysis SOCI C3998 -- Organizing Innovation SOCI C3108 -- Demographic Homophilies

Peter Bearman