Sam Ventura
36-315
Today: Networks

Department of Statistics
Carnegie Mellon University

March 29, 2017

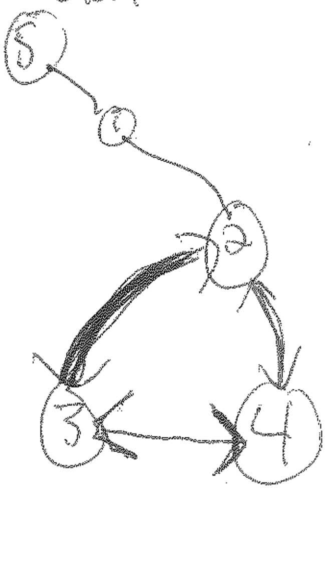
In degrees: "how many people follow me on Twitter"
Out degrees: "how many people do I follow ---"

Networks

cities, airports, train stations, PAT bus stop
wikipedia
actors, actresses, directors, movies, etc
family tree
→ animals?

Examples of networks:

research / co-authorship of economists
social networks: FB, twitter, IG, etc
electrical circuits
human brain / neural networks
internet



Basic Network Structure:

Nodes: "observations" — can be people, servers, actors, neurons, cities, airports, etc

Edges / links: connection between nodes in the network
→ or type of connection

undirected, ↓ LinkedIn, FB pre 2009ish
directed, ↓ twitter, IG
weighted → strength of the link/connection
↓ FB now/newsfeeds

path: distances between nodes — only care minimum/shortest

Clique: "cluster of nodes"

Adjacency is opposite of distance

$A_{ij} = g(D_{ij}) = g(d_{ij})$, where g is a monotonic decreasing function

Adjacency Matrices

Adjacency Matrix: representation of the network in matrix form

$$\underset{n \times n}{A} = \begin{array}{c|c} & 1 \quad 2 \quad \cdots \quad n \\ \hline 1 & \\ 2 & A_{ij} \text{ is edge weight / "link strength" /} \\ \vdots & \text{presence or absence of a link} \\ n & \text{between nodes } i \,\&\, j \end{array}$$

$A_{ij} \in \{0, 1\}$, $A_{ij} \neq A_{ji}$ } twitter

$A_{ij} \in [0, 1]$ } probability of a link
↳ FB's suggested friends

$A_{ij} \in [0, \infty)$

3/6

A_{ij} = link, represents the directed connection between nodes i and j

$A_{ij} \neq A_{ji}$ (doesn't have to be same, but can be)

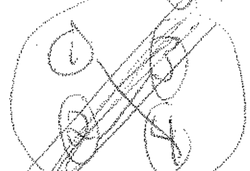
$A_{ij} = A_{ji} \Rightarrow$ undirected

Visualizing Networks

Layout: Trying to position the nodes to "maximize" visualization corresponding to specific criteria

force-directed algorithm: positions nodes such that

1) Edges between nodes are approximately the length
2) there are few crossing edges



Kamada-Kawai (smaller networks) → good for well-connected networks

poor choice for sparsely-connected network

4/6

Fructerman-Reingold: good for very large networks

adds extra emphasis on having even distribution of vertices