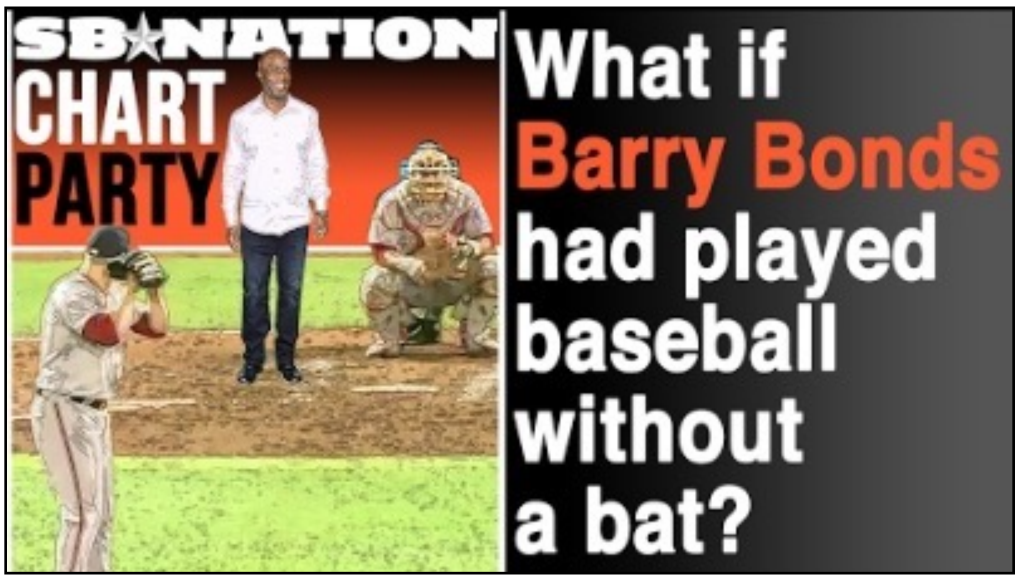
NETWORK ANALYSIS

CS 448B | Fall 2025

MANEESH AGRAWALA

1

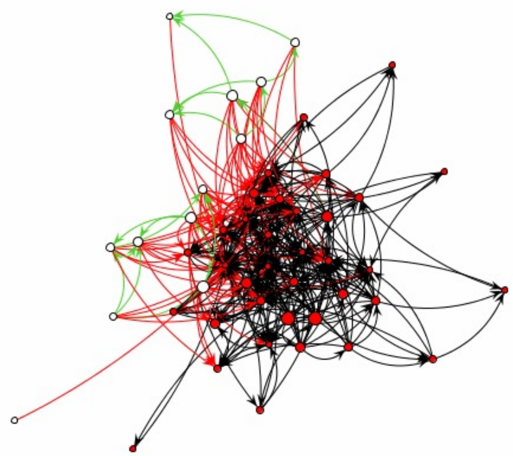


2

LAST TIME: NETWORK VISUALIZATION

5

LIMITATIONS OF NODE-LINK LAYOUTS



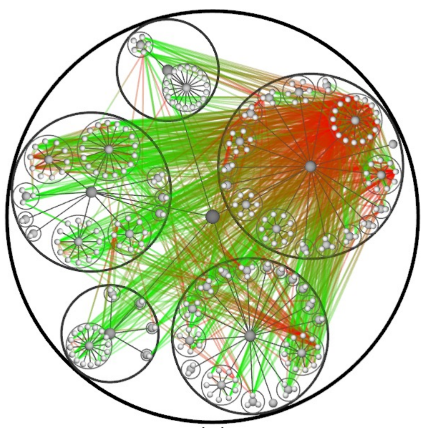
Edge crossings and occlusions!
Poor scalability…

6

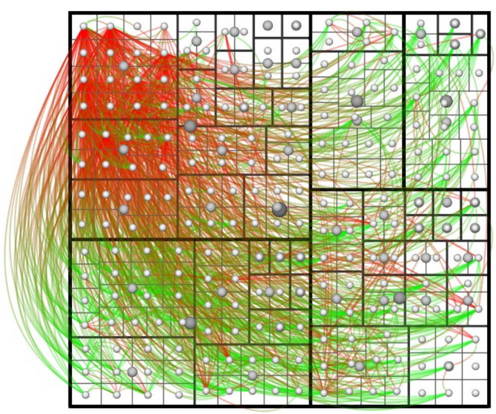
HIERARCHICAL EDGE BUNDLING

7

TREES WITH ADDITIONAL EDGES



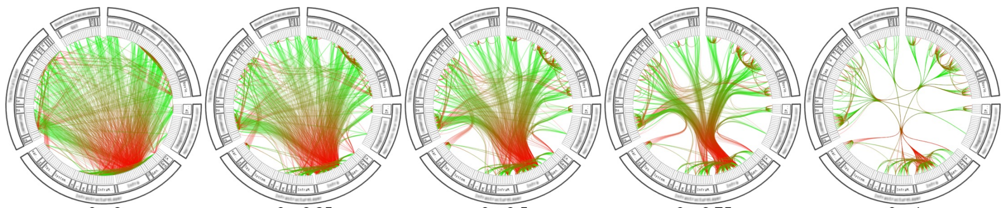
(a)



(b)

8

HIERARCHICAL EDGE BUNDLING

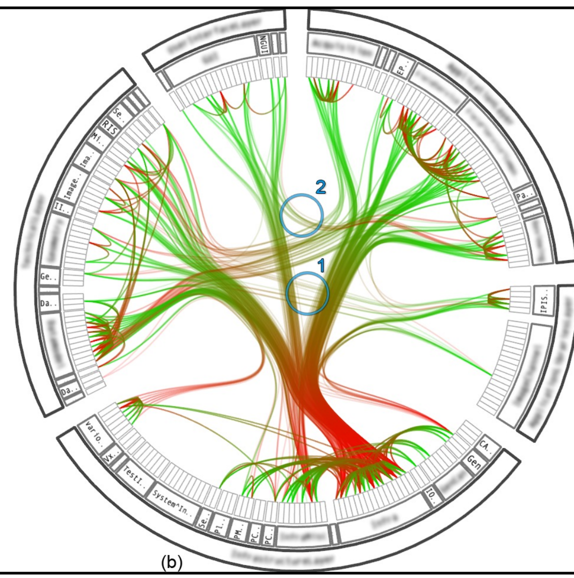


$\beta = 0$ $\beta = 0.25$ $\beta = 0.5$ $\beta = 0.75$ $\beta = 1$

Given a tree with additional *adjacency* edges (usually between leaves)

Bundle edges with varying amounts of tension – helping to reveal common connections between subtrees

9

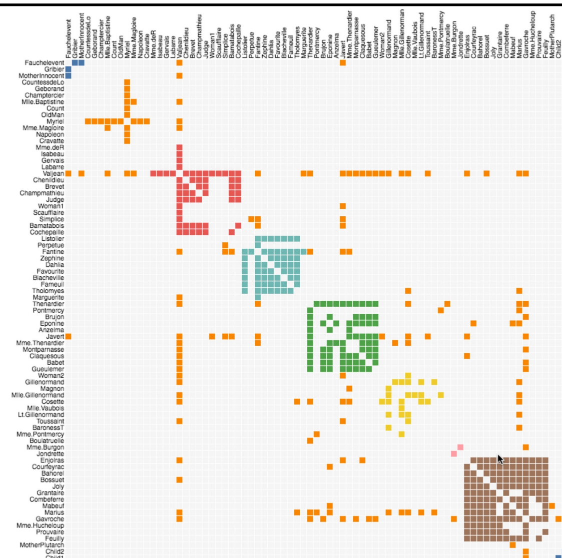


(b)

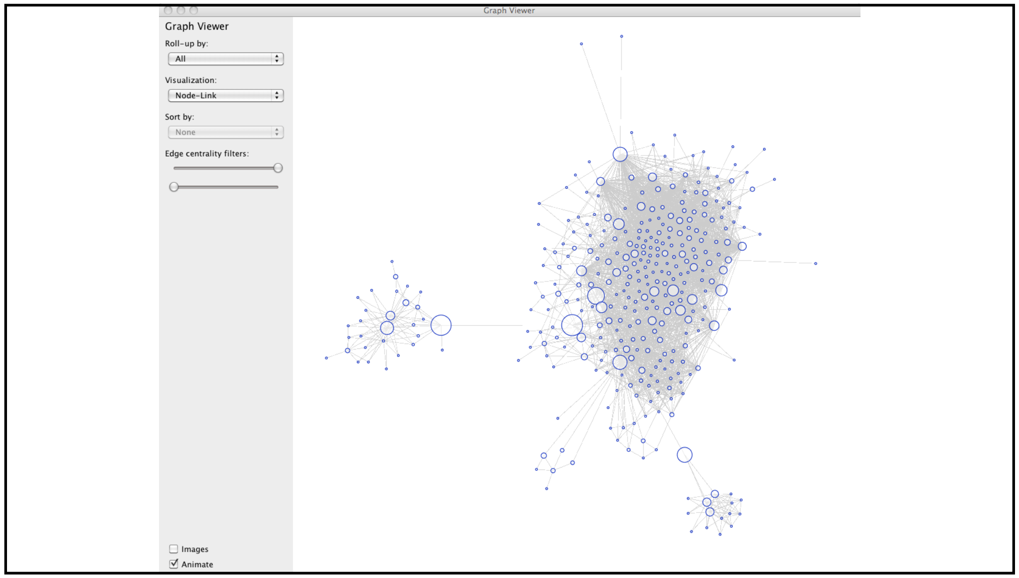
10

MATRIX DIAGRAMS

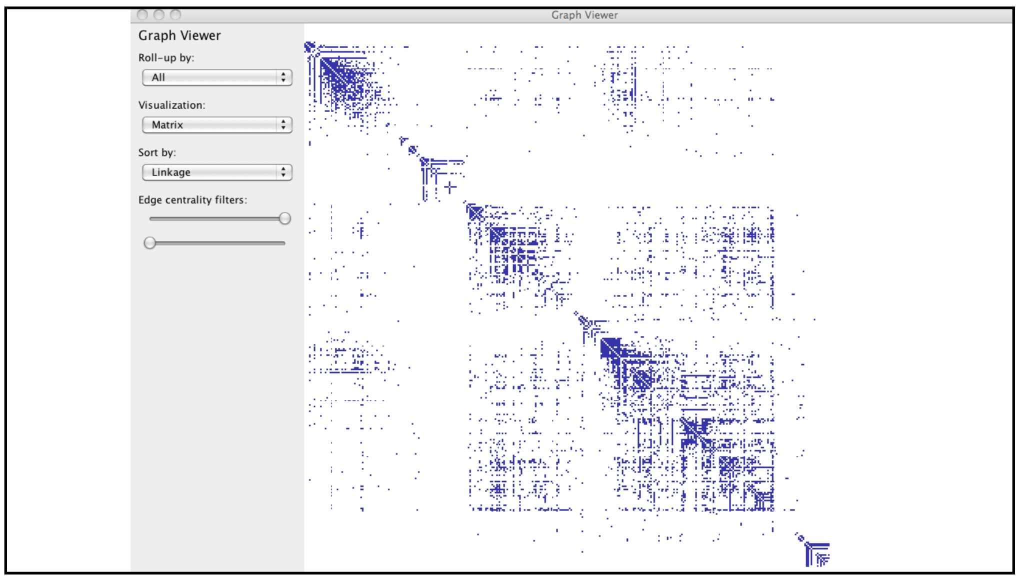
15



16


Graph Viewer
Graph Viewer
Roll-up by:
All
Visualization:
Node-Link
Sort by:
None
Edge centrality filters:
Images
Animate

17


Graph Viewer
Graph Viewer
Roll-up by:
All
Visualization:
Matrix
Sort by:
Linkage
Edge centrality filters:

18

SUMMARY: TREES AND NETWORKS

Tree Layout

Indented / Node-Link / Enclosure / Layers

Focus+Context techniques for scale

Graph Layout

Sugiyama Layout

Force-Directed Layout

Attribute-Driven Layout

Constraint Layout

Arc Diagrams

Matrix Diagrams

20

ANNOUNCEMENTS

21

FINAL PROJECT

Design Reviews Dec 1 and Dec 3

Data analysis/explainer

Analyze dataset in depth & make a visual explainer

Deliverables

An article with multiple different interactive visualizations

Short video (2 min) demoing and explaining the project

Schedule

Project proposal: Today!

Design Review and Feedback: 10th week of quarter, 12/1 and 12/3

Final code and video: Sun 12/7 8pm

Grading

Groups of up to 3 people, graded individually

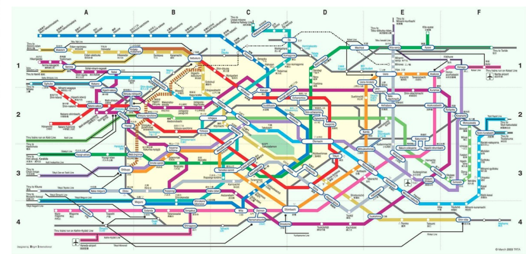
Clearly report responsibilities of each member

22

NETWORK ANALYSIS

25

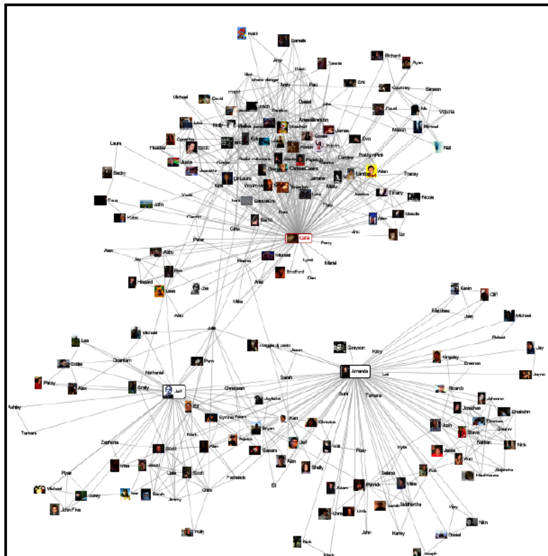
TRANSPORTATION



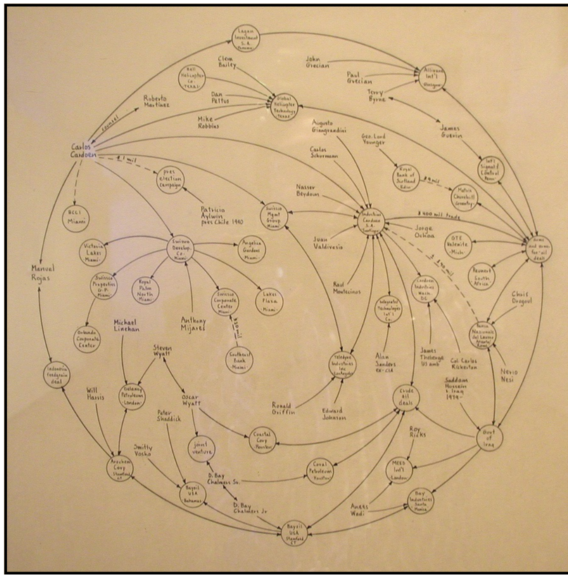
http://www.lx97.com/maps/

26

SOCIAL NETWORKS



28



"Mark Lombardi (1951-2000) draws on the major political and financial scandals of the day to create large-scale linear diagrams that at first glance look like celestial maps; a closer reading reveals the intricate web of connections that lurk beneath current headlines.

Lombardi, M. 'George W. Bush, Harken Energy and Jackson Stephens, ca 1979–90'

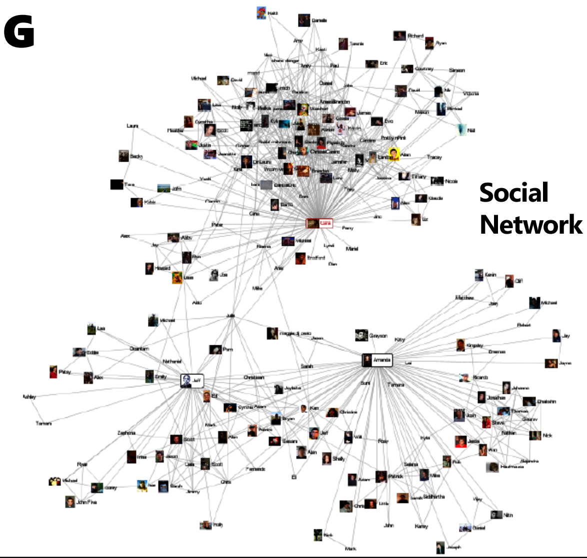
29

ACTORS & MOVIES (BIPARTITE)

30

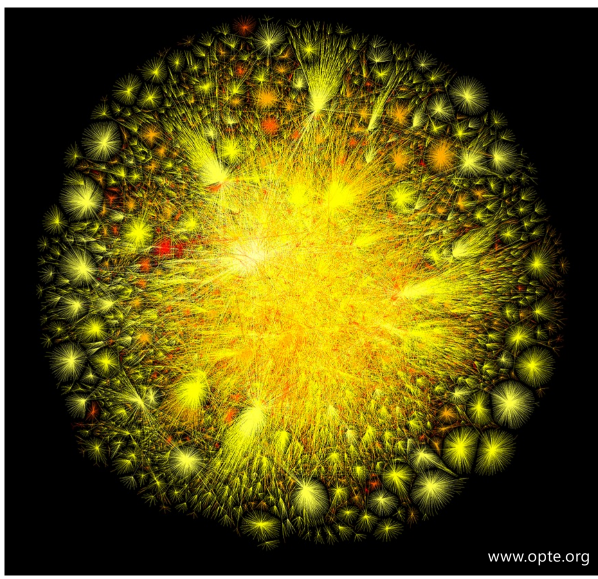
CHARACTERIZING NETWORKS

What does it look like?



Social Network

33



Size?
Density?
Centrality?
Clustering?
Components?
Cliques?
Motifs?
Avg. path length?
…

34

NETWORK ANALYSIS TASKS [Pretorius 2013]

Structure-based: relationships and connectivity

Attribute-based: properties associated with node or link

Browsing: follow paths in the data

Estimation: summarization and temporal changes

35

NETWORK ANALYSIS TASKS [Pretorius 2013]

Structure-based: relationships and connectivity

- Find all the friends of friends of Kermit
- Find all the people who are friends of Animal and Gonzo
- Find shortest path between two people: Six degrees of separation

Attribute-based: properties associated with node or link

Browsing: follow paths in the data

Estimation: summarization and temporal changes

36

NETWORK ANALYSIS TASKS [Pretorius 2013]

Structure-based: relationships and connectivity
- Find all the friends of friends of Kermit
- Find all the people who are friends of Animal and Gonzo
- Find shortest path between two people: Six degrees of separation

Attribute-based: properties associated with node or link
- Find all friends of Fozzie that are students at Stanford (node property)
- Find all friends of Fozzie that are their family (link property)

Browsing: follow paths in the data

Estimation: summarization and temporal changes

37

NETWORK ANALYSIS TASKS [Pretorius 2013]

Structure-based: relationships and connectivity
- Find all the friends of friends of Kermit
- Find all the people who are friends of Animal and Gonzo
- Find shortest path between two people: Six degrees of separation

Attribute-based: properties associated with node or link
- Find all friends of Fozzie that are students at Stanford (node property)
- Find all friends of Fozzie that are their family (link property)

Browsing: follow paths in the data
- Find Kermit's friend with first name Beaker and then find Beaker's mentor Bunsen

Estimation: summarization and temporal changes

38

NETWORK ANALYSIS TASKS [Pretorius 2013]

Structure-based: relationships and connectivity

- Find all the friends of friends of Kermit
- Find all the people who are friends of Animal and Gonzo
- Find shortest path between two people: Six degrees of separation

Attribute-based: properties associated with node or link

- Find all friends of Fozzie that are students at Stanford (node property)
- Find all friends of Fozzie that are their family (link property)

Browsing: follow paths in the data

- Find Kermit's friend with first name Beaker and then find Beaker's mentor Bunsen

Estimation: summarization and temporal changes

- How does Miss Piggy's friend group change over the course of a year

39

TODAY

Learning Objectives

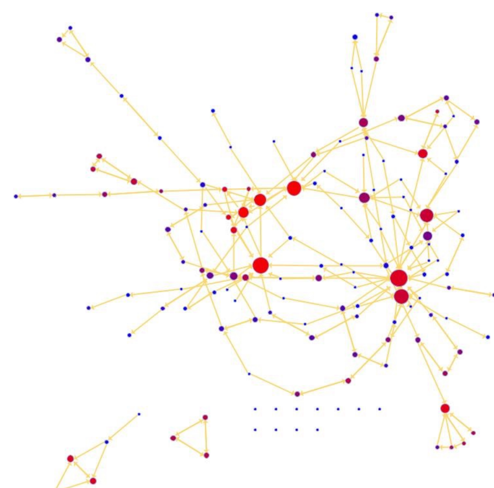
1. Measures of importance/centrality
2. Extracting community structure
3. Simulating network models

40

CENTRALITY

41

HOW FAR APART ARE THINGS?



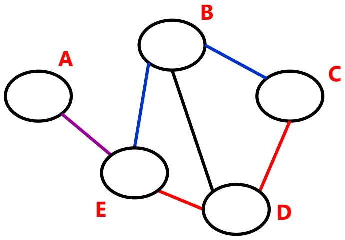
42

DISTANCE: SHORTEST PATHS

Shortest path (geodesic path)

The shortest sequence of links connecting two nodes
Not always unique

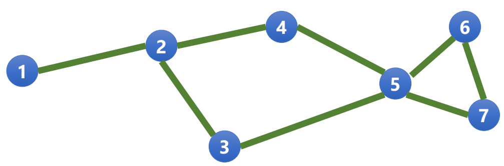
A and C are connected by 2 shortest paths
A – E – B - C
A – E – D - C



43

DISTANCE: SHORTEST PATHS

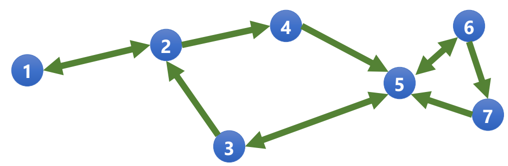
Shortest path from 2 to 3: 1



44

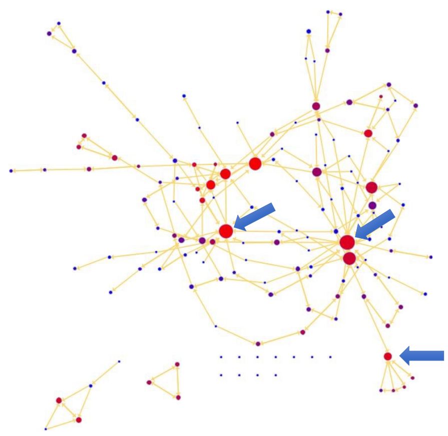
DISTANCE: SHORTEST PATHS

Shortest path from 2 to 3?



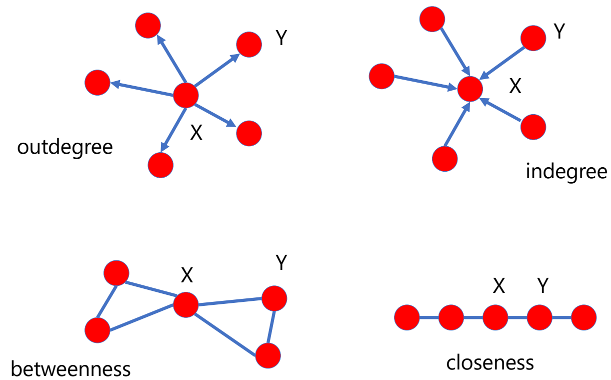
45

MOST IMPORTANT NODE



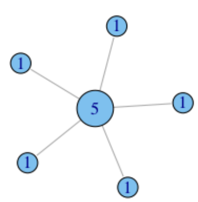
46

CENTRALITY



47

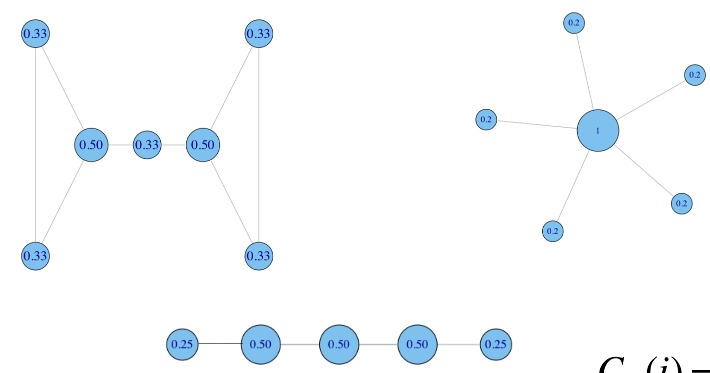
DEGREE CENTRALITY (UNDIRECTED)



$$C_D = d(i)$$

48

NORMALIZED DEGREE CENTRALITY



$$C_D(i) = \frac{d(i)}{N-1}$$

49

WHEN IS DEGREE NOT SUFFICIENT?

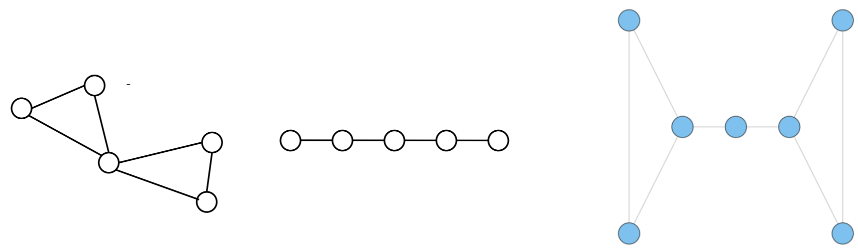
Does not capture

- Ability to broker between groups
- Likelihood that information originating anywhere in the network reaches you

50

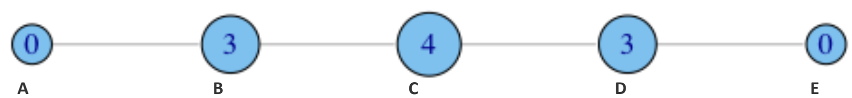
BETWEENESS

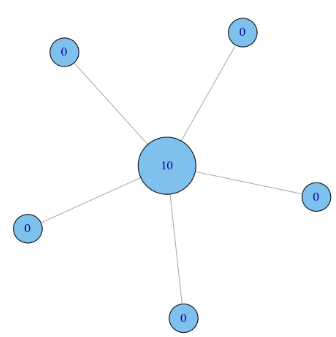
Assuming nodes communicate using the most direct (shortest) route, how many pairs of nodes have to pass information through target node?



51

BETWEENESS - EXAMPLES





52

BETWEENESS - DEFINITION

$$C_B(i) = \sum_{j,k \neq i, j<k} g_{jk}(i) / g_{jk}$$

g_{jk} = the number of shortest paths connecting *jk*
$g_{jk}(i)$ = the number of shortest paths containing *i*.

Normalization:

$$C'_B(i) = C_B(i) / [(n-1)(n-2)/2]$$

number of pairs of vertices excluding the vertex itself

53

WHEN ARE C_d, AND C_b NOT SUFFICIENT?

Does not capture

Likelihood that information originating anywhere in the network reaches you

55

CLOSENESS - DEFINITION

e.g., which node is closest to the *center* of the graph

Closeness Centrality:

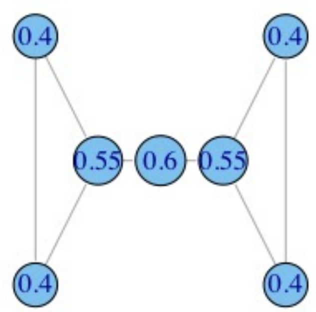
$$C_c(i) = \left[\sum_{j=1, j \neq i}^{N} d(i,j) \right]^{-1}$$

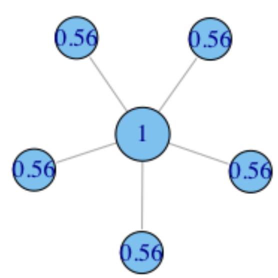
Normalized Closeness Centrality

$$C_C^{'}(i) = (C_C(i)) / (N-1) = \frac{N-1}{\sum_{j=1, j \neq i}^{N} d(i,j)}$$

56

CLOSENESS - EXAMPLES



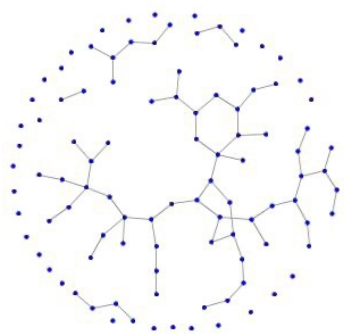


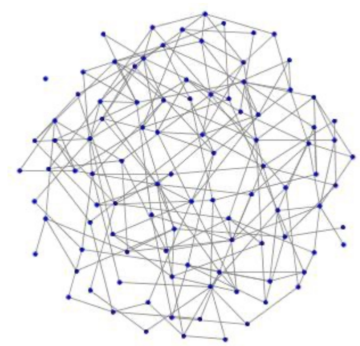
57

COMMUNITY STRUCTURE

67

HOW DENSE IS IT?





density = e / e_{max}

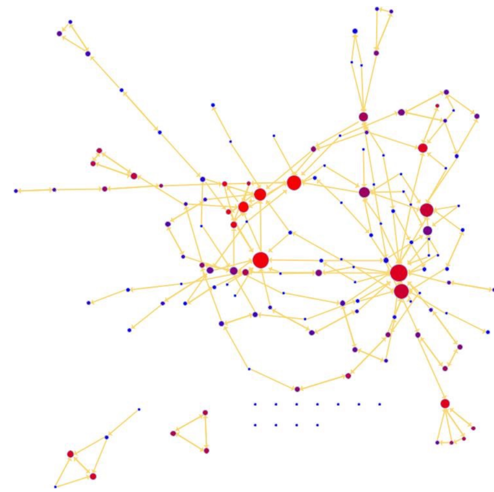
Max. possible edges:

Directed: $e_{max} = n*(n-1)$

Undirected: $e_{max} = n*(n-1)/2$

68

IS EVERYTHING CONNECTED?



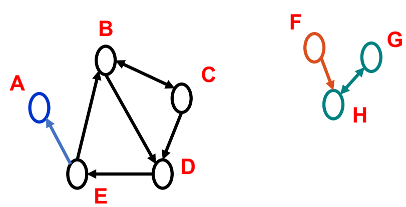
69

CONNECTED COMPONENTS - DIRECTED

Strongly connected components

Each node in component can be reached from every other node in component by following directed links

B C D E
A
G H
F



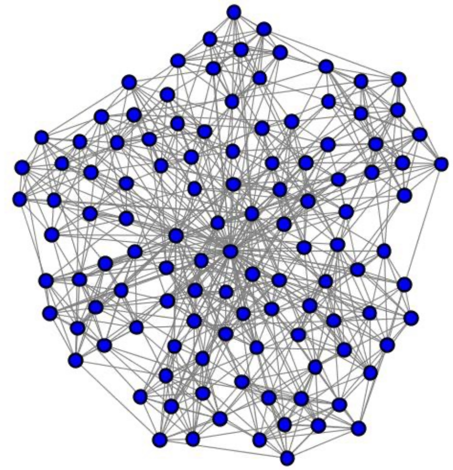
Weakly connected components

Each node can be reached from every other node by following links in either direction

A B C D E
G H F

70

COMMUNITY FINDING - CLUSTERING



73

HIERARCHICAL CLUSTERING

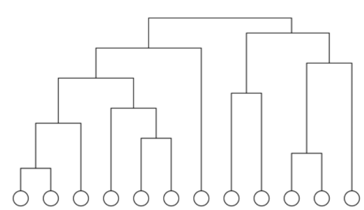
Process

Calculate affinity weights *W* for all pairs of vertices

Start: *N* disconnected vertices

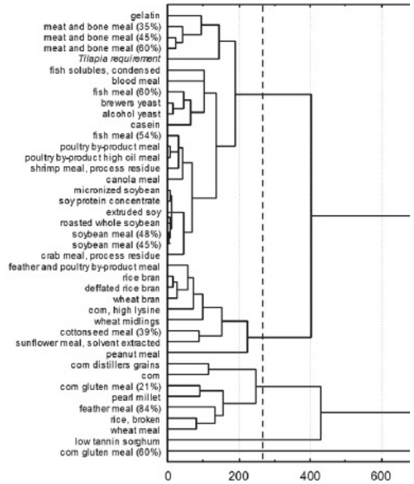
Add edges (one by one) between pairs of vertices/clusters in order of decreasing weight (use closest distance to compare clusters)

Result: nested components



74

CLUSTER DENDROGRAMS

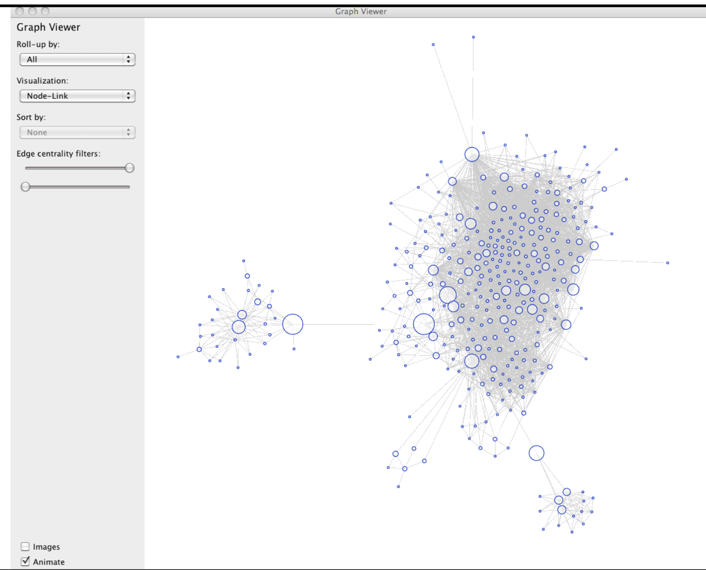


Depicts cluster trees produced by hierarchical clustering algorithms

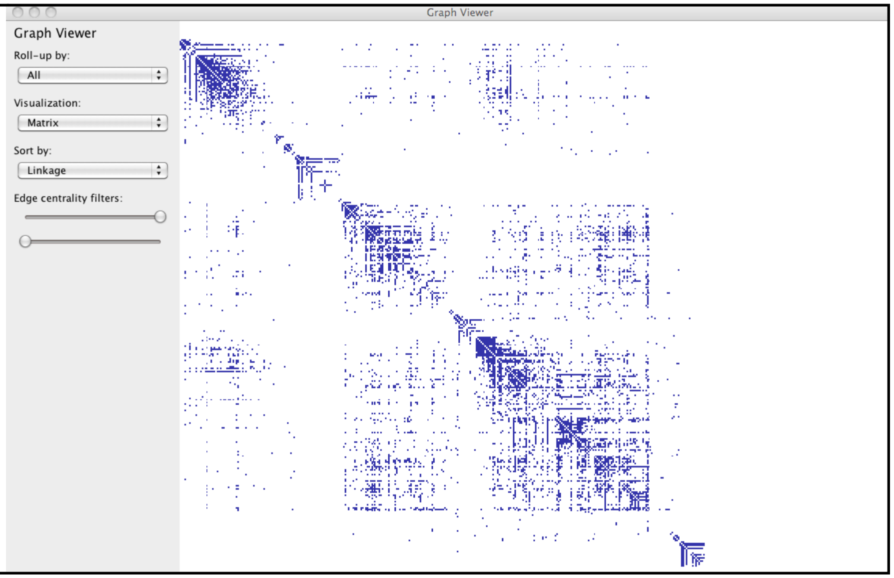
Leaf nodes arranged in line, internal node depth indicates order/value at which clusters merge

Basic recursive layout with orthogonal two-segment edges

75



76

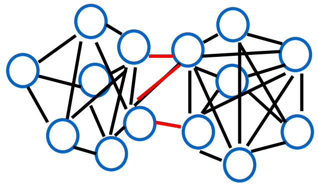


77

BETWEENESS CLUSTERING

Girvan and Newman 2002 iterative algorithm:

Compute C_b of all *edges*
Remove edge i where $C_b(i) == max(C_b)$
Recalculate betweenness



80

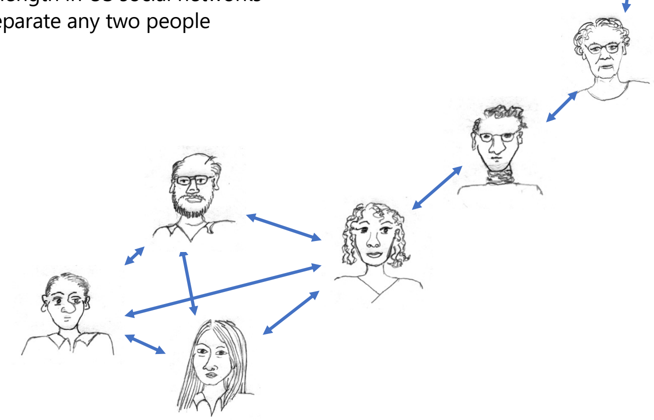
SIMULATING NETWORK MODELS

99

SMALL WORLD NETWORK

Milgram (1967)

Mean path length in US social networks
~ 6 hops separate any two people

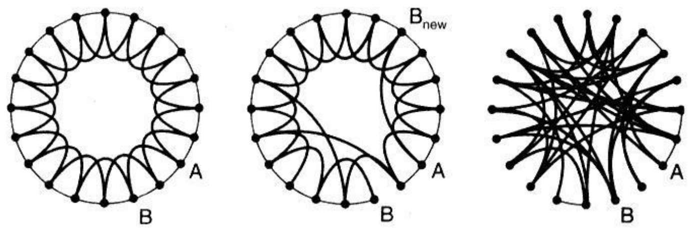


100

SMALL WORLD NETWORK

Watts and Strogatz 1998

a few random links in otherwise structured graph make network a small world



regular lattice:
my friend's friend is always my friend

small world:
mostly structured with a few random connections

random graph:
all connections random

101

DEFINING SMALL WORLD PHENOMENA

Properties

high clustering
low mean shortest path

$$C_{\text{network}} >> C_{\text{random graph}}$$

$$l_{\text{network}} \approx \ln(N)$$

Examples

neural network of C. elegans
semantic networks of languages
actor collaboration graph
food webs

102

SUMMARY

Structural analysis

- Centrality
- Community structure

Network analysis applicable in many domains

107