Nodes

- People
- Web pages
- Servers
- Articles
Edges

Undirected

Directed
Connectivity

**Connected component:**
subset of nodes where

— every node in the subset has a path to every other node

— that subset is not part of a larger set with that property
Small-world phenomenon


• 296 people asked to get a letter to a target near Boston by sending it to someone they knew on a first-name basis
Data

- Co-authorship networks
- Citation networks
- Social networks
- Hyperlink networks

https://snap.stanford.edu/data/
Adjacency Matrix

From:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Adjacency Matrix

\[ A_{3,1} = 1 \]
Degree (centrality)
Degree (centrality)

\[ \text{Degree}(3) = \sum_{i=1}^{5} A_{3,i} \]

\[ = A_{3,1} + A_{3,2} + A_{3,3} + A_{3,4} + A_{3,5} \]

From:

To:
Degree(i) = \sum_{j} A_{i,j}

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Degree

<table>
<thead>
<tr>
<th></th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>
Degree distribution

Degree | Number of nodes |
-------|----------------|
1      | 1              |
2      | 4              |
3      | 2              |
4      | 1              |
5      | 1              |
(Directed) Adjacency Matrix

From:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

To:

Under what circumstances is degree important?
Centrality

- Eigenvector centrality
  
  \[ \text{centrality}(i) = \sum_j \left[ A_{i,j} \times \text{centrality}(j) \right] \]

- Katz centrality
  
  \[ \text{centrality}(i) = \alpha \times \sum_j \left[ A_{i,j} \times \text{centrality}(j) \right] + \beta \]

- PageRank
  
  \[ \text{centrality}(i) = \alpha \times \sum_j \left[ A_{i,j} \times \frac{\text{centrality}(j)}{\text{outdegree}(j)} \right] + \beta \]
Geodesic path

Shortest path between two nodes
Closeness centrality

\[
\text{centrality}(i) = \frac{\sum_{j} \text{shortest}_\text{path}(i, j)}{n}
\]
Betweenness centrality

\[ \text{betweenness}(i) = \sum_{s,t} 1 \{ i \text{ is on the path from } s \text{ to } t \} \]
## Summary: centrality

<table>
<thead>
<tr>
<th>What’s important?</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of friends</td>
<td>Degree centrality</td>
</tr>
<tr>
<td>Number or importance of friends</td>
<td>Eigenvector, Katz centrality; PageRank</td>
</tr>
<tr>
<td>Distance from others</td>
<td>Closeness centrality</td>
</tr>
<tr>
<td>Middleman</td>
<td>Betweenness centrality</td>
</tr>
</tbody>
</table>
Tie strength

- “Strong” ties vs. “weak” ties
Tie strength


Maintained: click on news feed story/visit profile 3+ times

One-way: any directed message

Reciprocal: reciprocated message
Triadic closure

Two people (A and B) have a friend (C) in common; A and B are likely to become friends.

More likely the stronger the tie is between A-C and B-C.
Triadic closure

• Why?

• A and B have more opportunity to interact if both are friends with the same person

• A and B may trust each other if they’re both friends with the same person

• C has a matchmaking incentive
Structural balance
Structural balance
Structural balance
Clustering coefficient

- Probability that two randomly selected friends of A will be friends with each other
structural bridges

- early access to information
- ability to combine different sources of information
- gatekeeper between components
Assortativity
Assortativity

\[
\frac{1}{2} \sum_{i,j} \left[ A_{i,j} \times I\{\text{if } \text{node}(i) = \text{node}(j)\} \right]
\]

\[
-\frac{1}{2} \sum_{i,j} \left[ \frac{\text{outdegree}(i) \times \text{outdegree}(j)}{2m} \right] \times I\{\text{if } \text{node}(i) = \text{node}(j)\}
\]

\(m = \text{total number of edges in network}\)
Assortativity

Project presentation

Monday April 25 (6) + Wednesday April 27 (5)

10 min presentation +
3-5 min questions
YOUR CONFERENCE PRESENTATION

HOW YOU PLANNED IT:

INTRODUCE YOURSELF

DESCRIBE OUTLINE OF TALK

MOTIVATION

METHODOLOGY AND EXPERIMENT DESIGN

RESULTS

CONCLUSIONS

APPLAUSE

ENGAGING Q&A

15 MINUTES
Final report

• 8 pages, single spaced.

• Complete description of work undertaken
  • Data collection
  • Methods
  • Experimental details
  • Comparison with past work
  • Analysis

• See many of the papers we’ve read this semester for examples.
Final report

• **Clarity.** For the reasonably well-prepared reader, is it clear what was done and why? Is the paper well-written and well-structured?

• **Originality.** How original is the approach or problem presented in this paper? Does this paper break new ground in topic, methodology, or content? How exciting and innovative is the research it describes?

• **Soundness.** Is the technical approach sound and well-chosen? Second, can one trust the claims of the paper -- are they supported by proper experiments, proofs, or other argumentation?

• **Substance.** Does this paper have enough substance, or would it benefit from more ideas or results? Do the authors identify potential limitations of their work?

• **Evaluation.** To what extent has the application or tool been tested and evaluated? Does this paper present a compelling argument for

• **Meaningful comparison.** Do the authors make clear where the presented system sits with respect to existing literature? Are the references adequate? Are the benefits of the system/application well-supported and are the limitations identified?

• **Impact.** How significant is the work described? Will novel aspects of the system result in other researchers adopting the approach in their own work?