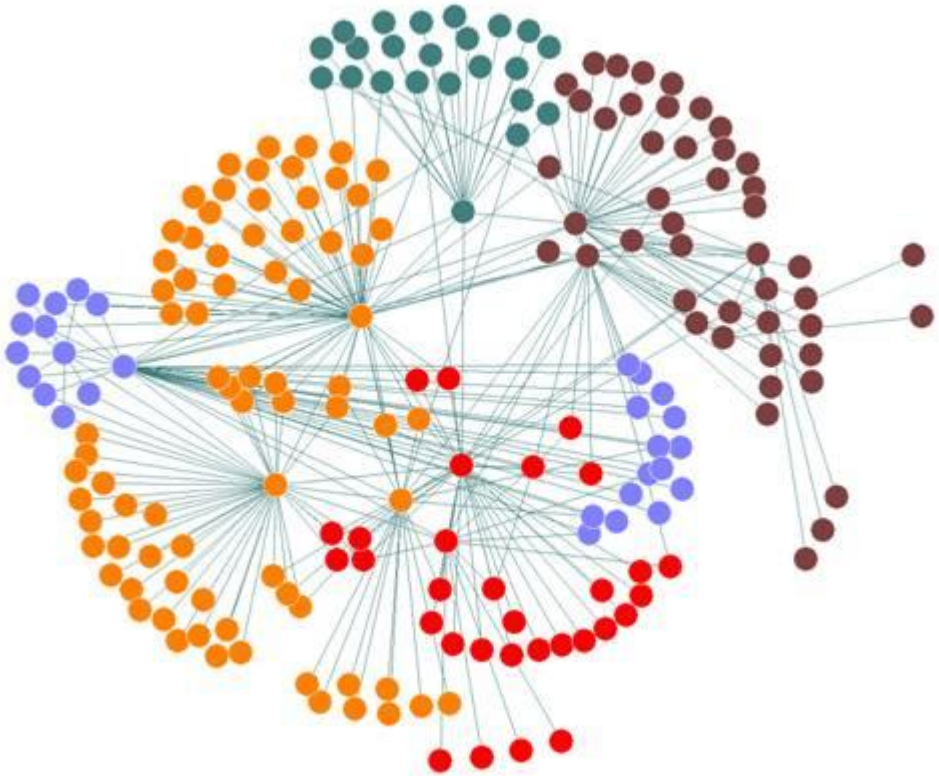
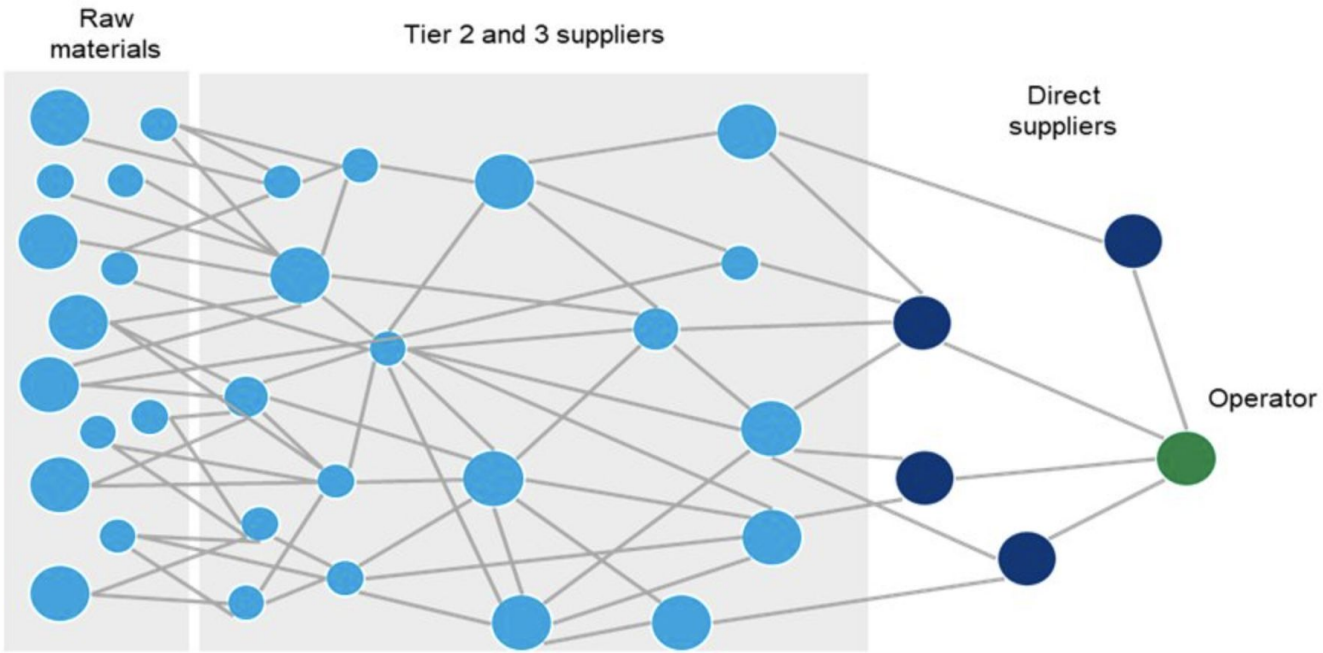
MIDS W207
Applied Machine Learning

Week 12
Live Session Slides

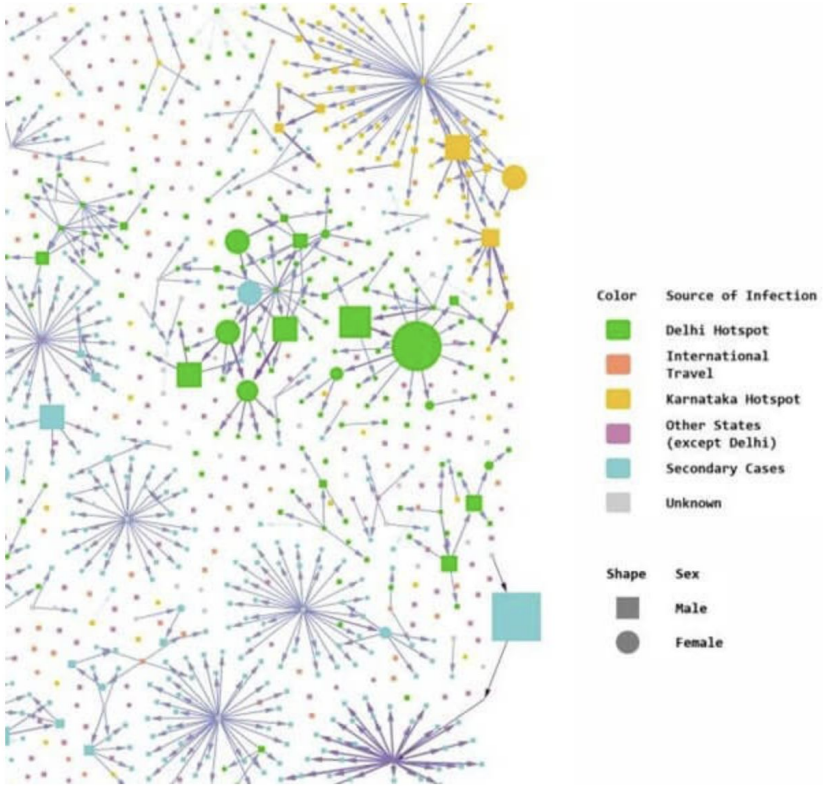
Social Network Analysis



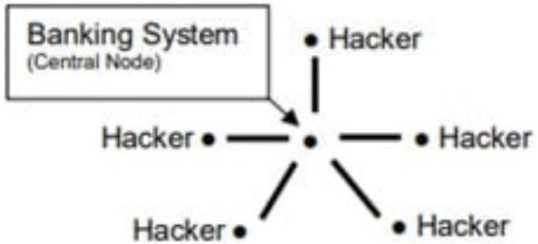
Social Network Analysis: Applications



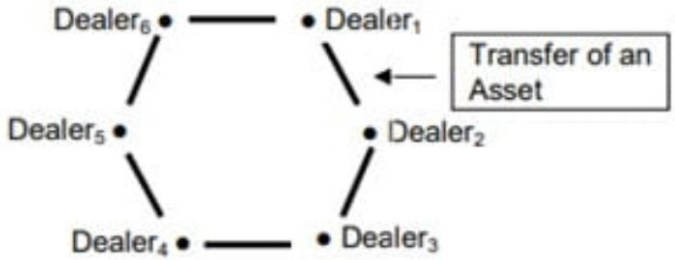
Social Network Analysis: Applications



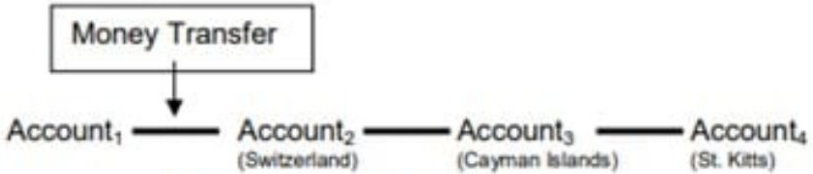
Social Network Analysis: Applications



Denial of Service-Hacker Attack (Star)



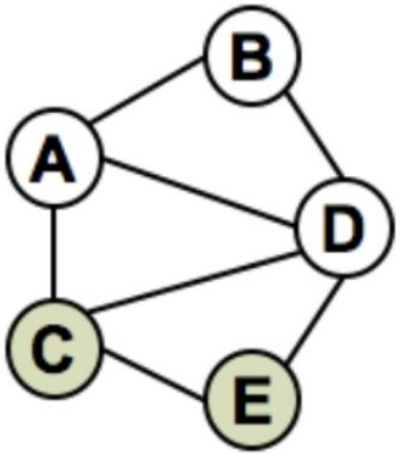
Networking Fraud Ring (Circle)



Money Laundering (Chain)

Network Representation

Nodes



Edges

Real World Networks

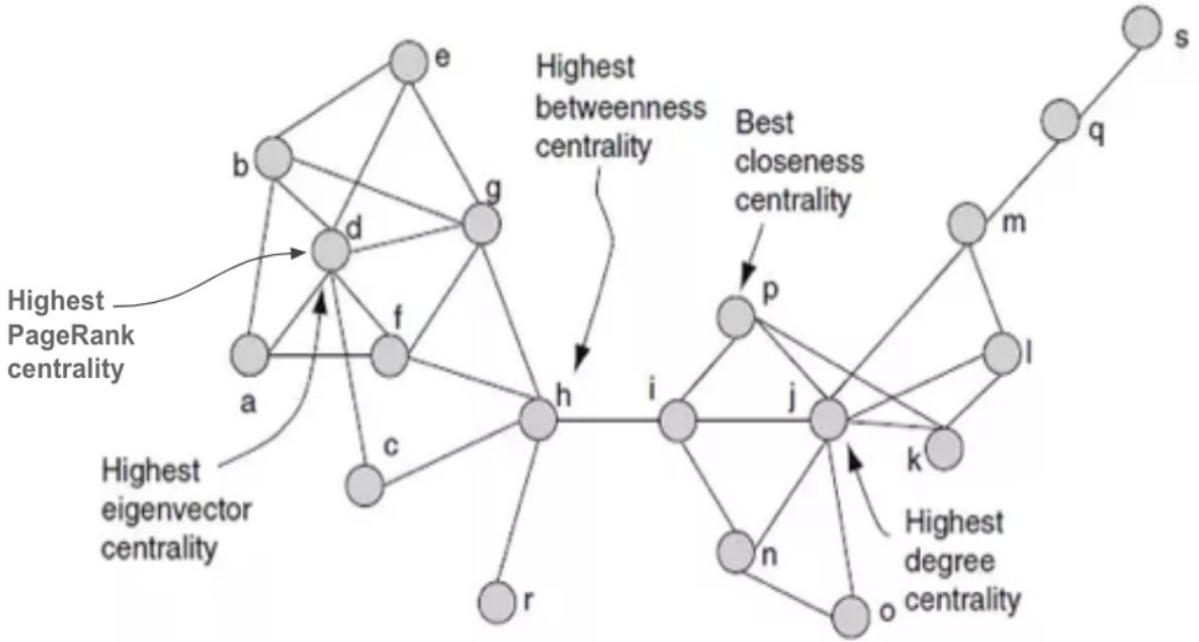
Centrality Measures

Degree

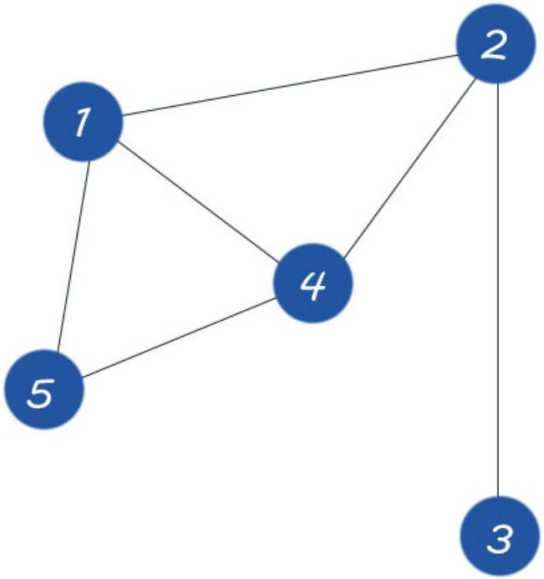
Eigenvector/ PageRank

Closeness

Betweenness



Adjacency Matrix



$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

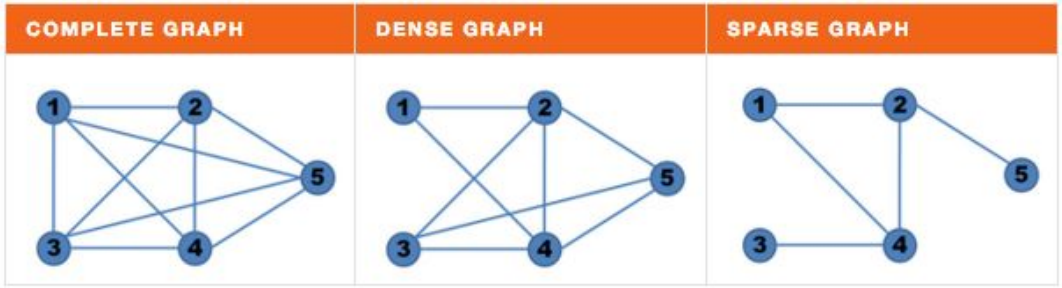
Graph Level Properties

Density

Connected Components

Degree Distribution

Giant Components



Graph Level Properties

Density

Connected Components

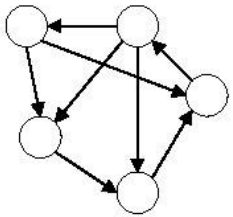
Degree Distribution

Giant Components

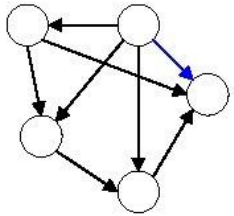
Strongly-Connected

Graph G is ***strongly connected*** if, for every u and v in V, there is some path from u to v and some path from v to u.

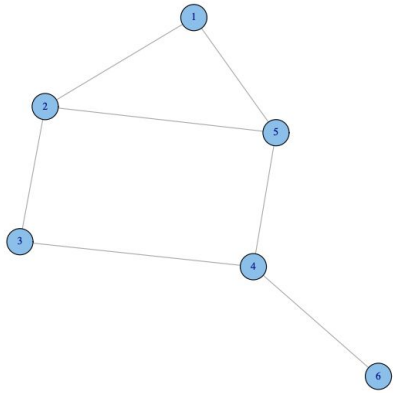
Strongly Connected



Not Strongly Connected



Graph Level Properties



Density

Connected Components

Degree Distribution

Giant Components

$$\langle K \rangle = \frac{\sum_{i=1}^{n} deg(i)}{N} = \frac{2L}{N}$$

node	degree
1	2
2	3
3	2
4	3
5	3
6	1

degree	frequency
1	1/6
2	2/6
3	3/6

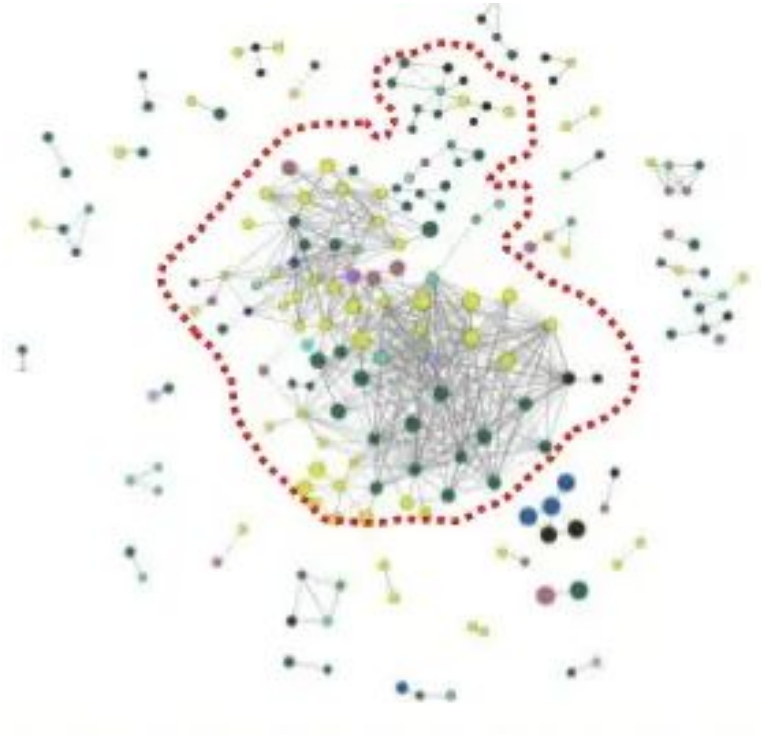
Graph Level Properties

Density

Connected Components

Degree Distribution

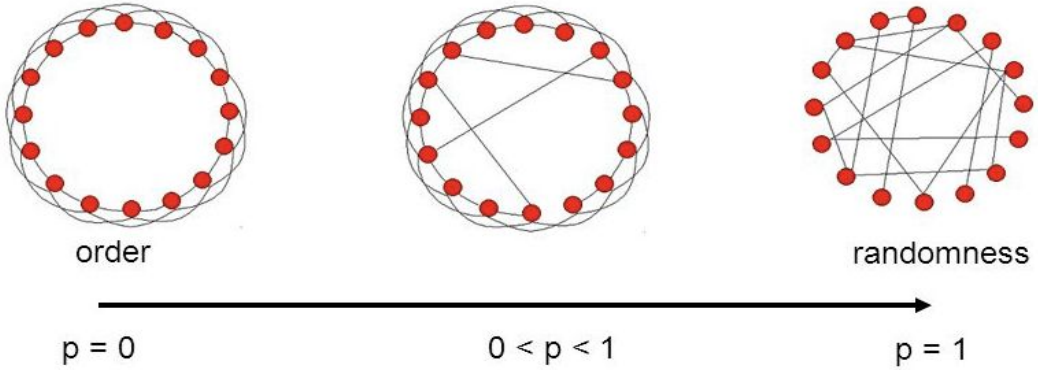
Giant Components



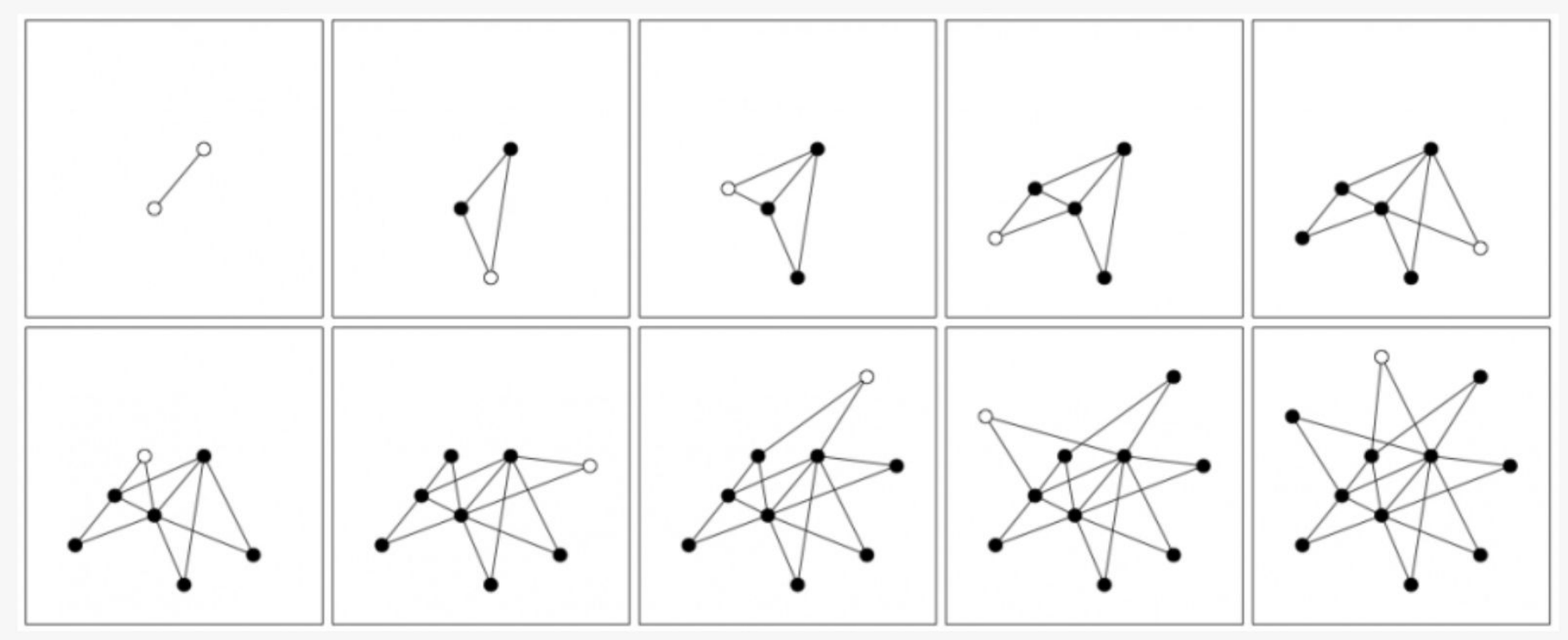
Watts strogatz model

Watts and Strogatz model [WS98]

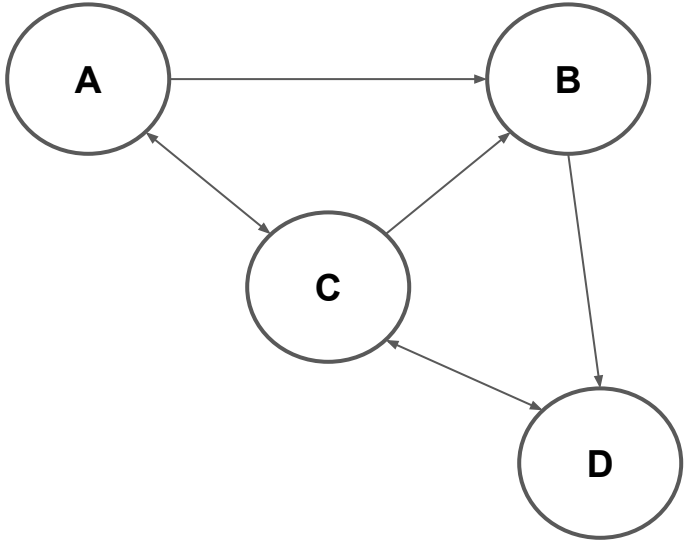
- Start with a ring, where every node is connected to the next k nodes
- With probability p, rewire every edge (or, add a shortcut) to a uniformly chosen destination.
  - Granovetter, "The strength of weak ties"



Barabasi Albert model



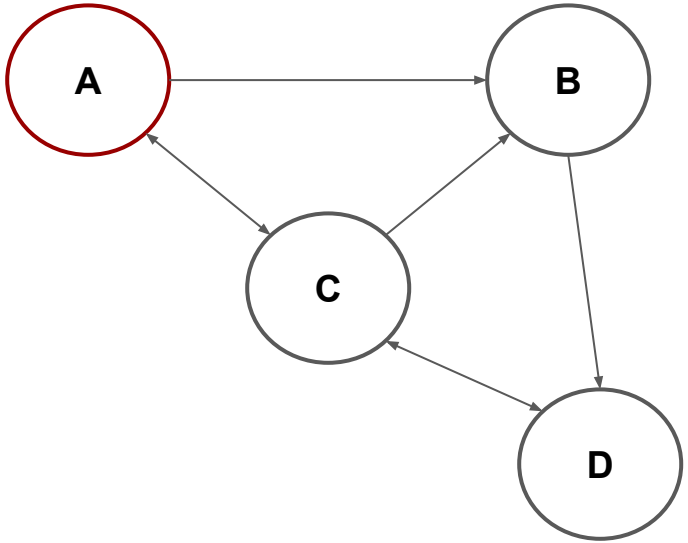
Page Rank Algorithm



	Iteration 1	Iteration 2	Iteration 3	Page Rank
A	1/4			
B	1/4			
C	1/4			
D	1/4			

$$PR_{t+1}(P_i) = \Sigma\ [PR_t(P_j)] / C(P_j)$$

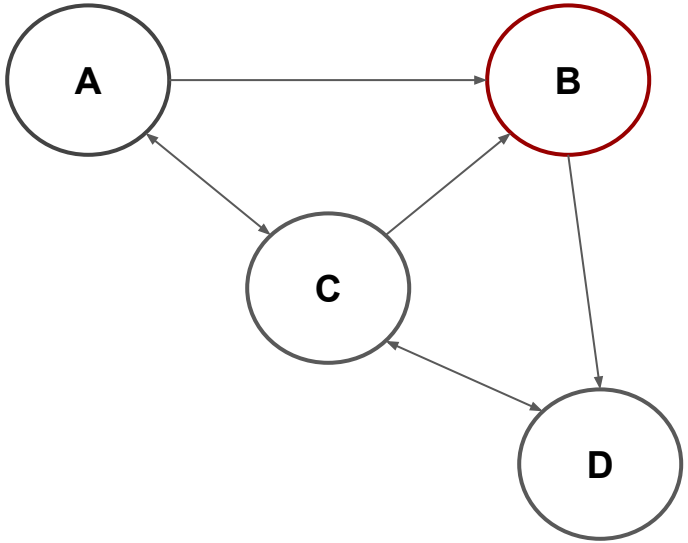
Page Rank Algorithm



PR(A)= (¼) / 3

	Iteration 1	Iteration 2	Iteration 3	Page Rank
A	1/4	1/12		
B	1/4			
C	1/4			
D	1/4			

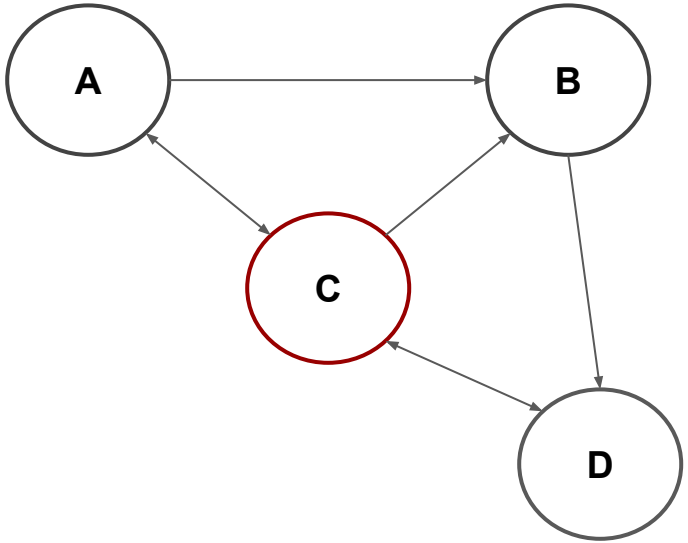
Page Rank Algorithm



PR(B)= ((¼) / 2) + ((¼)/3)

	Iteration 1	Iteration 2	Iteration 3	Page Rank
A	1/4	1/12		
B	1/4	2.5/12		
C	1/4			
D	1/4			

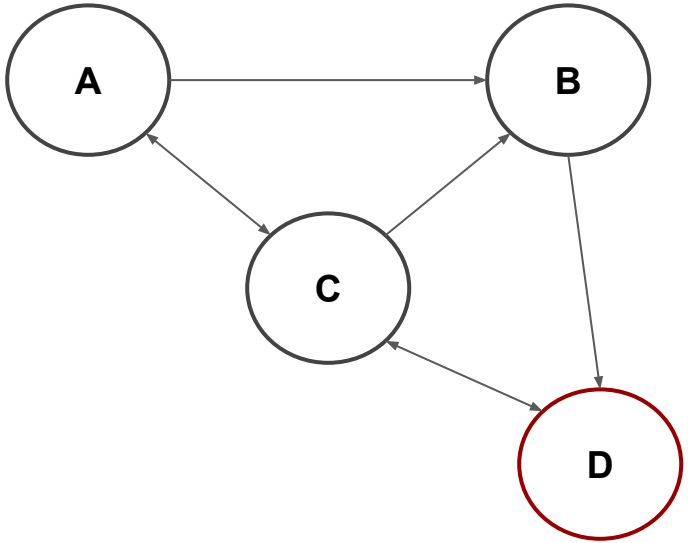
Page Rank Algorithm



	Iteration 1	Iteration 2	Iteration 3	Page Rank
A	1/4	1/12		
B	1/4	2.5/12		
C	1/4	4.5/12		
D	1/4			

PR(C)= ((¼) / 2) + ((¼)/1)

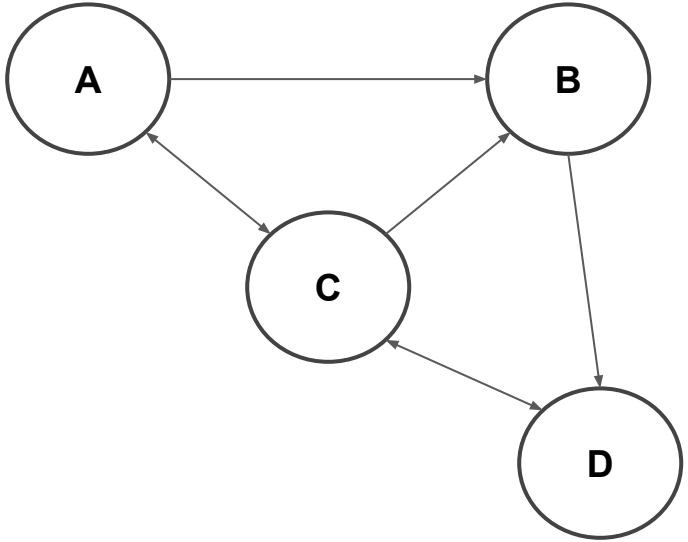
Page Rank Algorithm



	Iteration 1	Iteration 2	Iteration 3	Page Rank
A	1/4	1/12		
B	1/4	2.5/12		
C	1/4	4.5/12		
D	1/4	4/12		

PR(D)= ((¼) / 3) + ((¼)/1)

Page Rank Algorithm



	Iteration 1	Iteration 2	Iteration 3	Page Rank
A	1/4	1/12	1.5/12	1
B	1/4	2.5/12	2/12	2
C	1/4	4.5/12	4.5/12	4
D	1/4	4/12	4/12	3

$$PR_{t+1}(P_i) = \Sigma\ [PR_t(P_j)] / C(P_j)$$

Network Visualization

iGraph

NetworkX

Gephi

UCINET

NetLogo

Code Review