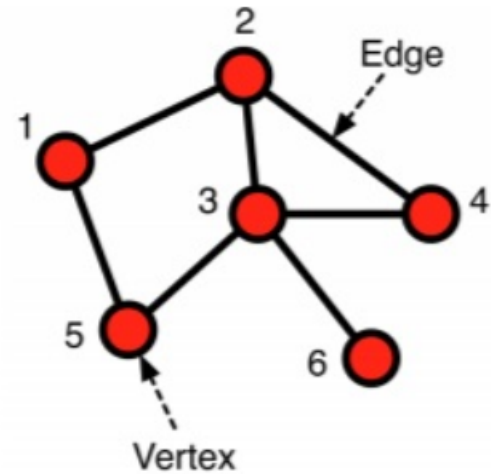# MUNI

Session 4 – Network Analysis

Oleg Deev

Undirected Graphs

– A graph with undirected edge interactions between variables is an undirected graph.

– These graphs produce a class of models commonly known as undirected graphical models, which are suitable for analyzing similarities and correlated behaviors
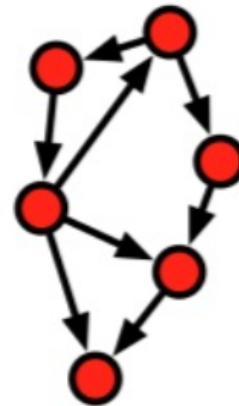
MUNI

Directed Graphs

– A graph with directed edges between variables is a directed graph

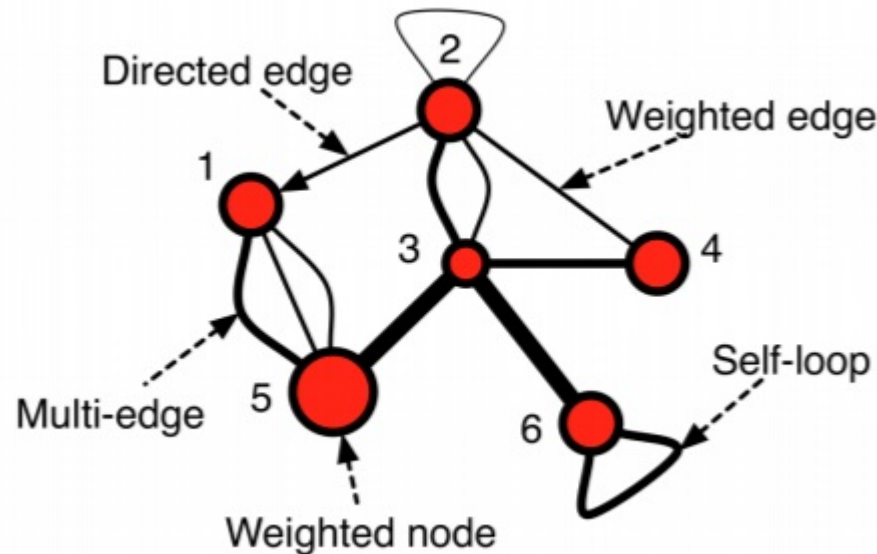– A directed graph without cycles is a directed acyclic graph (DAG)
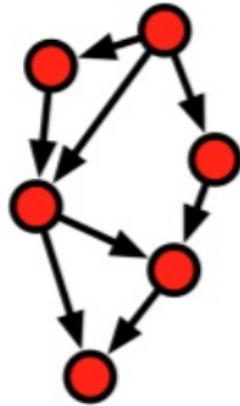
directed acyclic graph   directed graph

MUNI

Other Network Types

- A partially directed acyclic graph (or chain graph) is a type of DAG that allows bi-directed edges.
  e.g. A –> B –> C
- A weighted graph is one that has a numeric value (weights) associated with each edge (node)

Bipartite Graphs

— A bipartite graph is an undirected graph in which variables are categorized into two sets, such that nodes in one set can only interact with those in the other set, and no two nodes in the same set are connected

MUNI
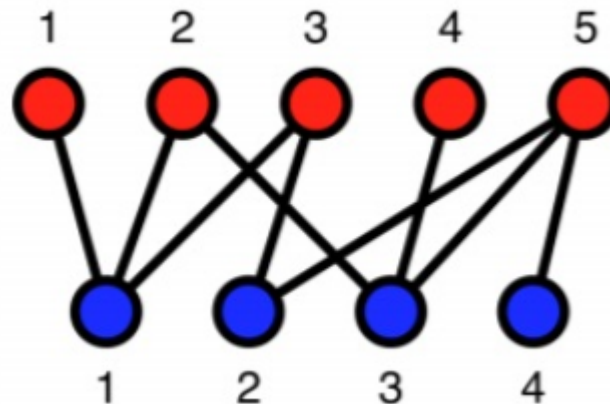
Multilayer Networks

— In reality, individual units belong to more than one network (in content, in space, in time)
— This give rise to multilayer networks

M U N I

Network Representation

- A graph-theoretic representation $G = (V; E; W)$ of relationships E (edges, links) of strengths W between units V (vertices, nodes)
- Nodes: individuals, companies, stocks, consumers, industries, countries
- Physical links:
  - import/export
  - borrowing/lending
  - common shareholdings
  - supply chains
- Statistical links:
  - correlations
  - similarities (commercial, financial, operations)

MUNI

Adjacency Matrix

$$A_{i,j} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connceted} \\ 0 & \text{otherwise} \end{cases}$$
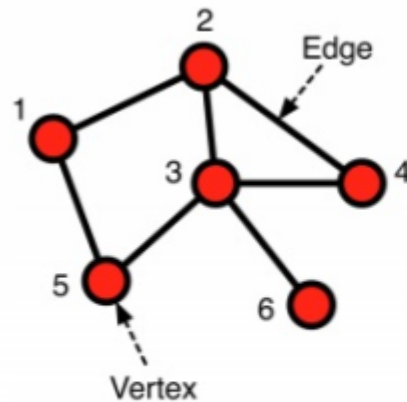
Example:
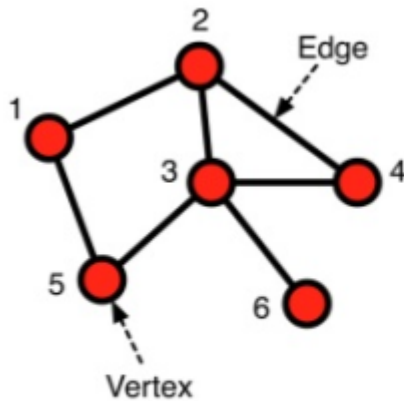
(a) Undirected Graph

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

(b) Adjacency Matrix

MUNI

Adjacency List

– The adjacency list stores only the non-zero elements of the adjacency matrix in a full list of all vertices.

Example:

1	2	5		
2	3	1	4	
3	2	5	4	6
4	2	3		
5	1	3		
6	3			

(a) Undirected Graph

(b) Adjacency List

MUNI

Edge List

– The edge list stores only the non-zero elements of the adjacency matrix.

Example:
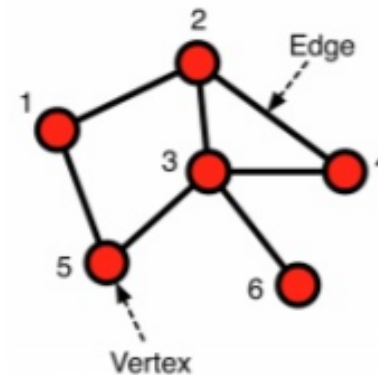
(a) Undirected Graph

$$\{(1,2),(1,5),(2,3),(2,4),(3,5),(3,6)\},$$

(b) Edge List

MUNI

Centrality Measures

— A common question we may ask when analyzing the structure of a network is which vertices are more important?

— To answer this question depends on what we mean by important. Define importance in terms
  — the most connected vertex
  — the vertex closest to all others
  — the vertex in the middle of the network
  — the most influential vertex

— The most common centrality measures
  — degree centrality
  — closeness centrality
  — betweenness centrality
  — eigenvector centrality

MUNI

Degree Centrality

– The simplest measure of importance is the degree centrality
– The vertex with largest degree exerts the greatest effect on the network
– Degree centrality: number of nearest neighbors

$$D(i) = \sum_{j=1}^{n} A_{ij} = \sum_{j=1}^{n} A_{ji}$$

– Sensitive to the addition of one more node
– We cannot compare degree centrality of vertices belonging to two different network
– Normalized degree centrality (relative metric to compare degree centrality)

$$N_{D(i)} = \frac{1}{n-1} D(i)$$

MUNI

Degree Centrality: Example
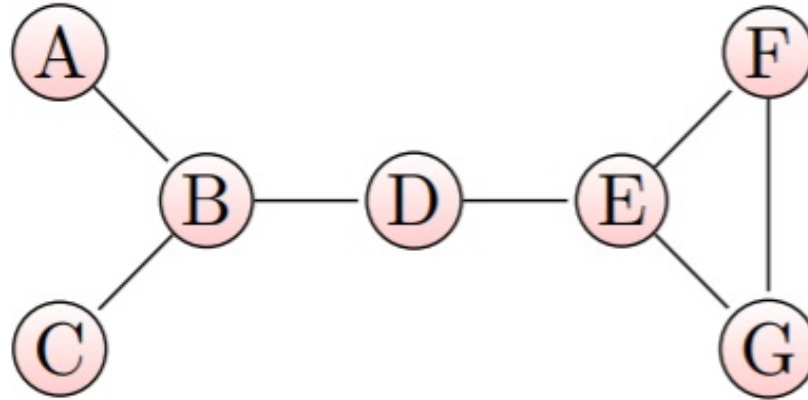
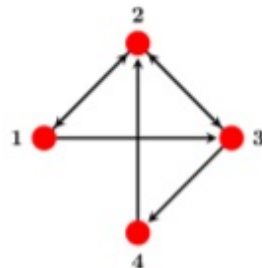| Node	Degree	Normalized
A	1	1/6
B	3	3/6
C	1	1/6
D	2	2/6
E	3	3/6
F	2	2/6
G	2	2/6

MUNI

In-Degree and Out-Degree

In-Degree: the number of edges incoming to a vertex

$$d_i^{\leftarrow} = \sum_{j=1}^{n} A_{ij}$$

Out-Degree: the number of edges outgoing from a vertex

$$d_i^{\rightarrow} = \sum_{j=1}^{n} A_{ij}$$

(a) Directed Graph

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

(b) Adjacency Matrix

node	out-degree	in-degree
1	2	1
2	2	3
3	2	2
4	1	1

(c) Degrees

MUNI

Closeness Centrality

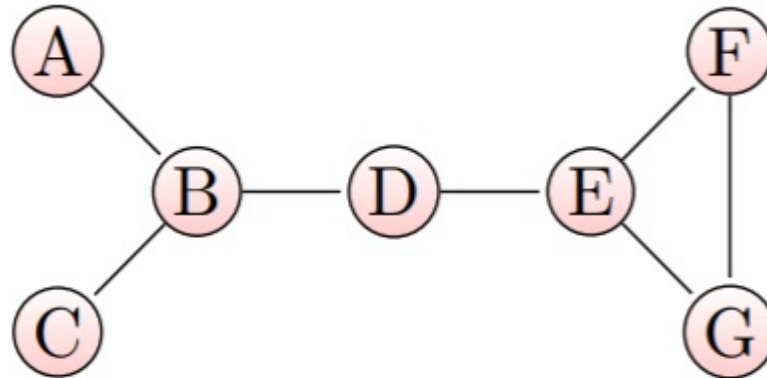— Measures how close a node is to all the other nodes in network (in terms of the shortest path)

$$C(i) = \frac{1}{\sum_{j=1}^{n} d(i,j)}$$

where $d(i,j)$ is the shortest path between the nodes i and j

— Normalized closeness centrality (relative metric to compare closeness centrality)

$$N_{c(i)} = (n-1)C(i)$$

MUNI

Closeness Centrality: Example

Node	Closeness	Normalized
A	1/16	6/16
B	1/11	6/11
C	1/16	6/16
D	1/10	6/10
E	1/11	6/11
F	1/15	6/15
G	1/15	6/15

MUNI

Betweenness Centrality

— The number of shortest paths from all vertices to all others that pass through a node.
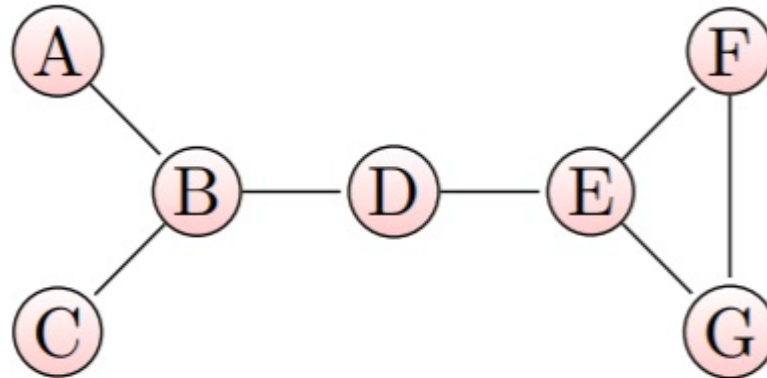
$$B(i) = \sum_{s \neq i \neq t} \frac{\delta_{st(i)}}{\delta_{st}}$$

where $st(i)$ is the number of shortest paths between s and t that pass through i, and st is the total number of shortest paths between s and t.

— Probability that a communication from s to t will go through i (the frequency of i lying on the shortest path)

— Normalized betweenness centrality

$$N_{B(i)} = \frac{2}{(n-1)(n-2)} B(i)$$

MUNI

Betweenness Centrality: Example

| Node	Betweeness	Normalized
A	0	0
B	9	9/15
C	0	0
D	9	9/15
E	8	8/15
F	0	0
G	0	0

MUNI

Eigenvector Centrality

– Eigenvector centrality measures assign an importance score to vertices in a way that is proportional to the importance scores of its neighbors hence the importance of a node depends on the importance of its neighbors

– Shows how a well-connected vertex is connected to other well-connected vertices

– This involves an eigenvector problem of the form

$$Av = \lambda_1 v$$

where A is the adjacency matrix, v is a vector containing the eigenvector centralities, and λ_1 is the largest eigenvalue of A

– PageRank is a form of eigenvector centrality

M U N I

Centrality Measures: Example
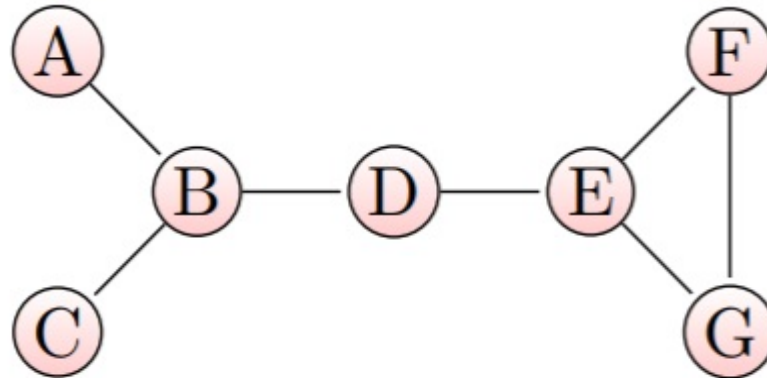
Node	Degree	Closeness	Betweeness	Eigenvector
A	0.1667	0.3750	0	0.2134
B	**0.5000**	0.5454	**0.6000**	0.4814
C	0.1667	0.3750	0	0.2136
D	0.3333	**0.6000**	**0.6000**	0.6574
E	**0.5000**	0.5454	0.5333	**1.0000**
F	0.3333	0.4000	0	0.7979
G	0.3333	0.4000	0	0.7979

MUNI

Centrality Measures: Example
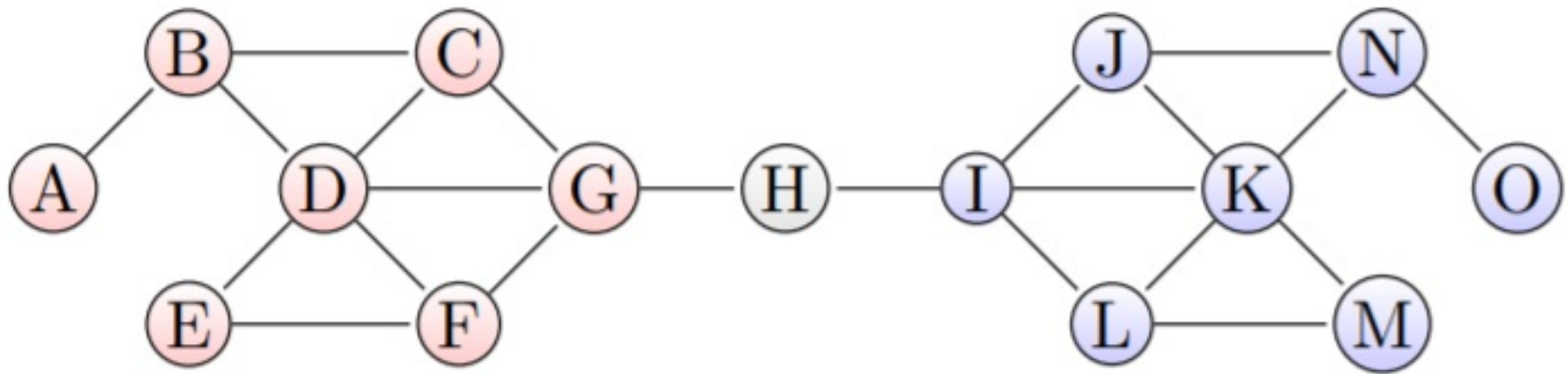
— Degree most central nodes: D & K
— Closeness most central nodes: H
— Betweenness most central nodes: H
— Eigenvector most central nodes: D & K

MUNI

Network Density

— The density of a network is the frequency of realized edges relative to potential edges

— Undirected Networks: For n number of nodes, there are $n(n - 1)/2$ possible edges

$$\text{den}(U_G) = \frac{|E_G|}{n(n-1)/2}$$

— Directed Networks: For n number of nodes, there are $n(n - 1)$ possible edges

$$\text{den}(D_G) = \frac{|E_G|}{n(n - 1)}$$

MUNI

Network Analysis

– Can we construct network, when we don't have information on the exact links?
  – Network reconstruction algorithms
  – Similarity networks
– Complexity reductions
  – Threshold networks
  – Spanning trees
– Simulations on networks
– Flow analysis
– Community/pattern identification

MUNI

Applications in Risk Management

– Market risk

   – Asset correlation networks
   – Volatility-based networks

– Liquidity and operational risks

   – Payment networks

– Counterparty and systemic risks

   – Exposure networks

M U N I