Social Network Analysis

Basic Concepts, Methods & Theory

University of Cologne
Johannes Putzke

Agenda

- Introduction
- Basic Concepts
- Mathematical Notation
- Network Statistics

Textbooks

- Hanneman & Riddle (2005) *Introduction to Social Network Methods,* available at *http*://faculty.ucr.edu/~hanneman/nettext/
- Wasserman & Faust (1994): *Social Network Analysis – Methods and Applications*, Cambridge: Cambridge University Press.

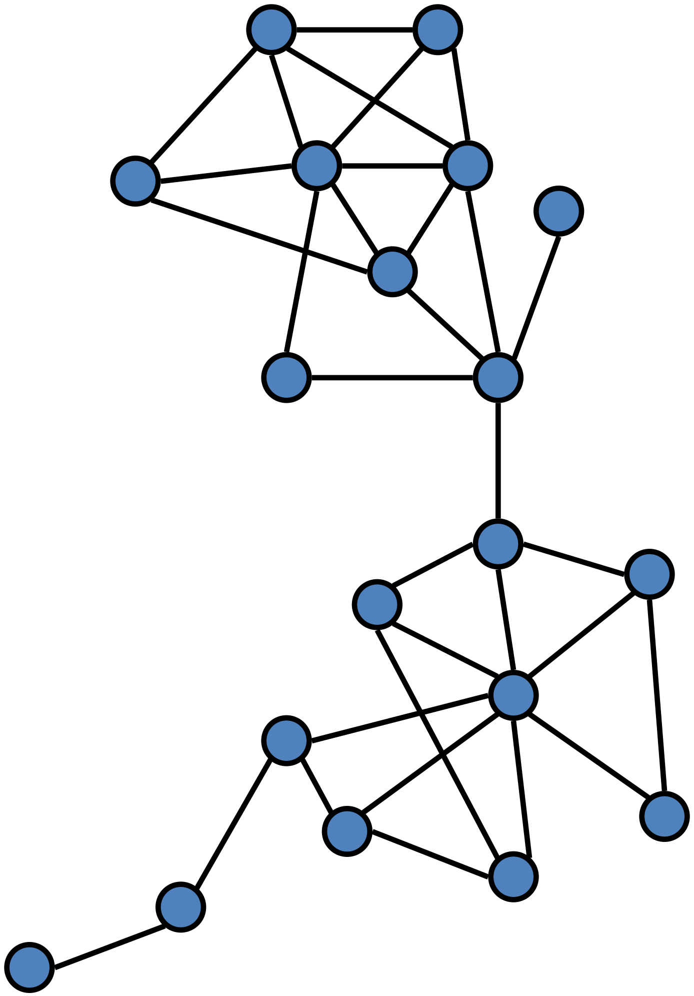
Introduction

Basic Concepts

What is a network?

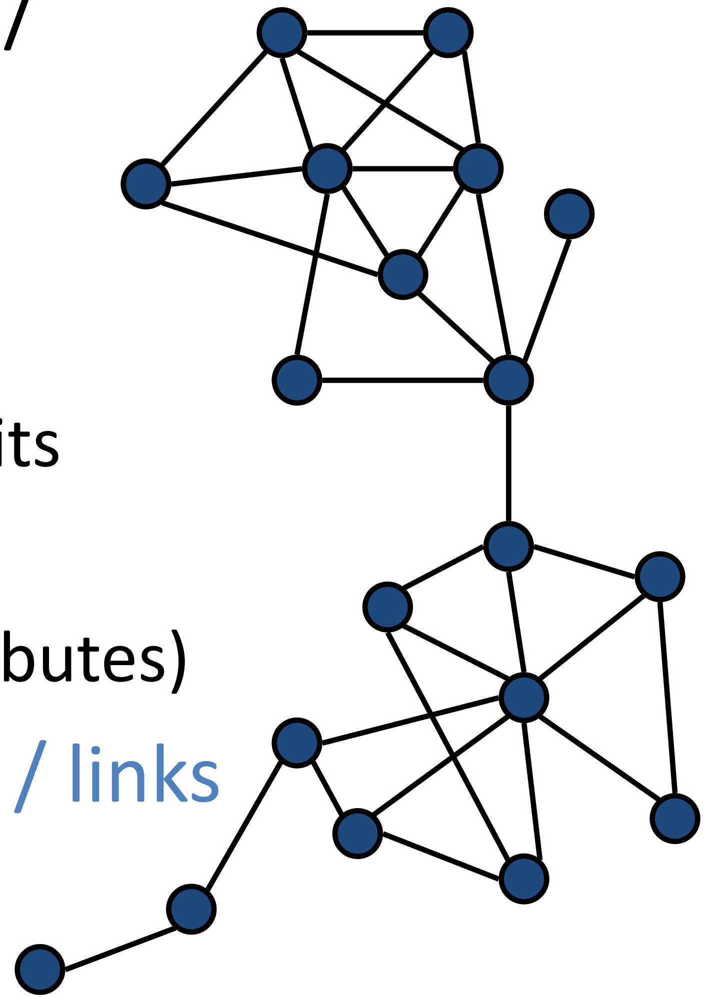
What is a Network?

- Actors / nodes / vertices / points
- Ties / edges / arcs / lines / links



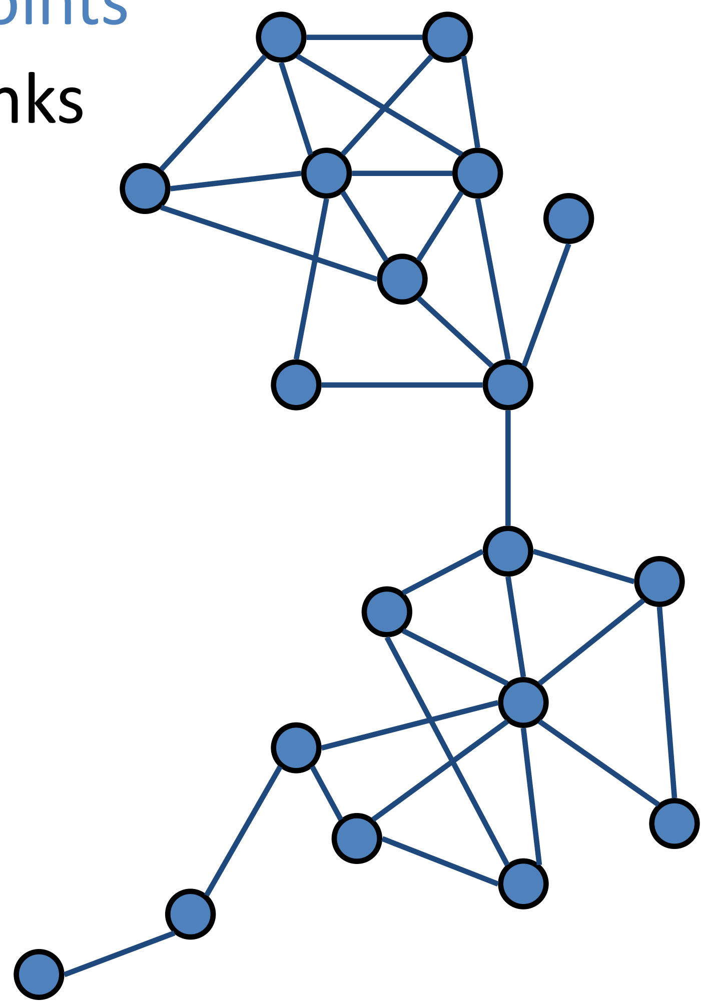
What is a Network?

- Actors / nodes / vertices / points
  - Computers / Telephones
  - Persons / Employees
  - Companies / Business Units
  - Articles / Books
  - Can have properties (attributes)
- Ties / edges / arcs / lines / links

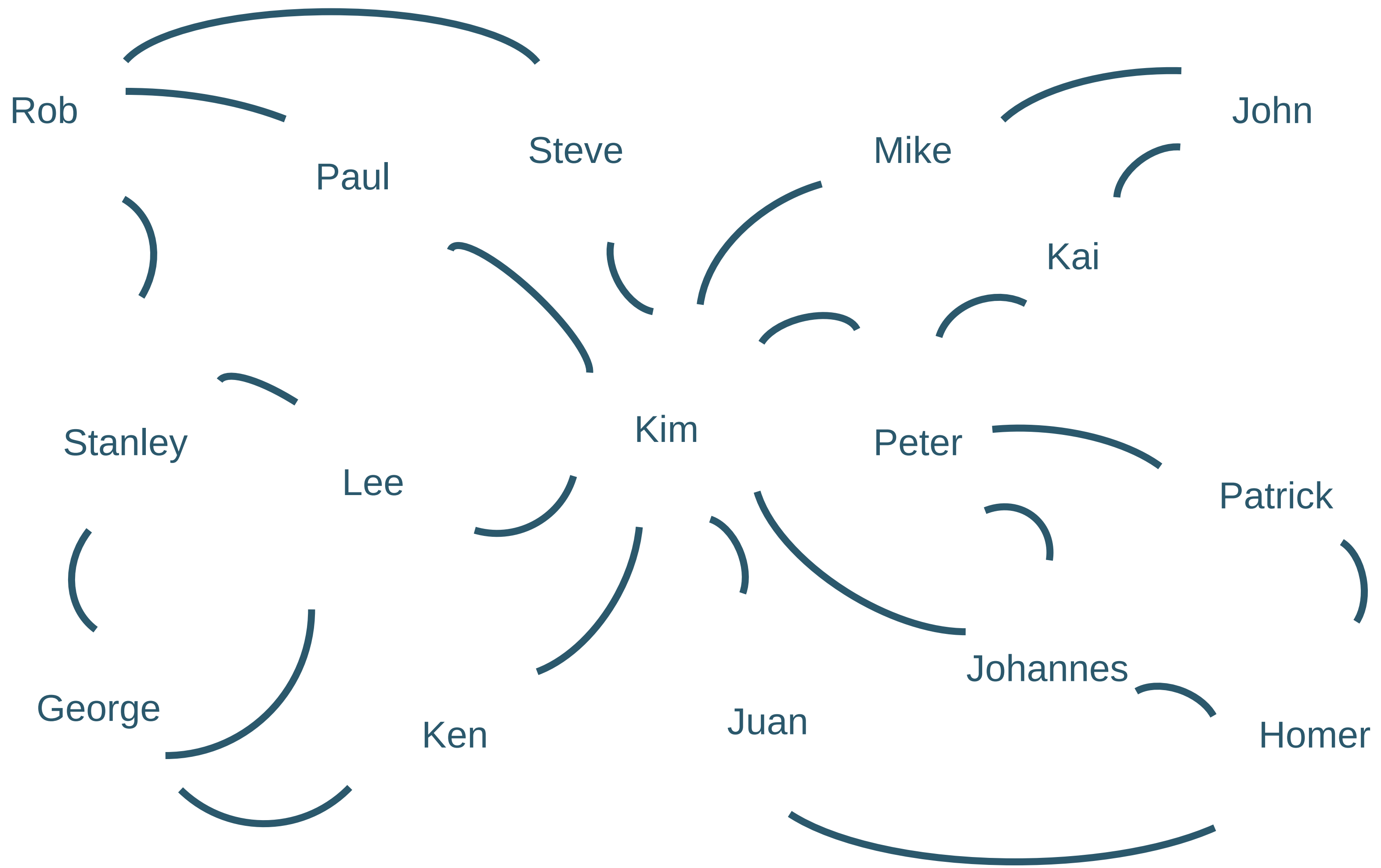


What is a Network?

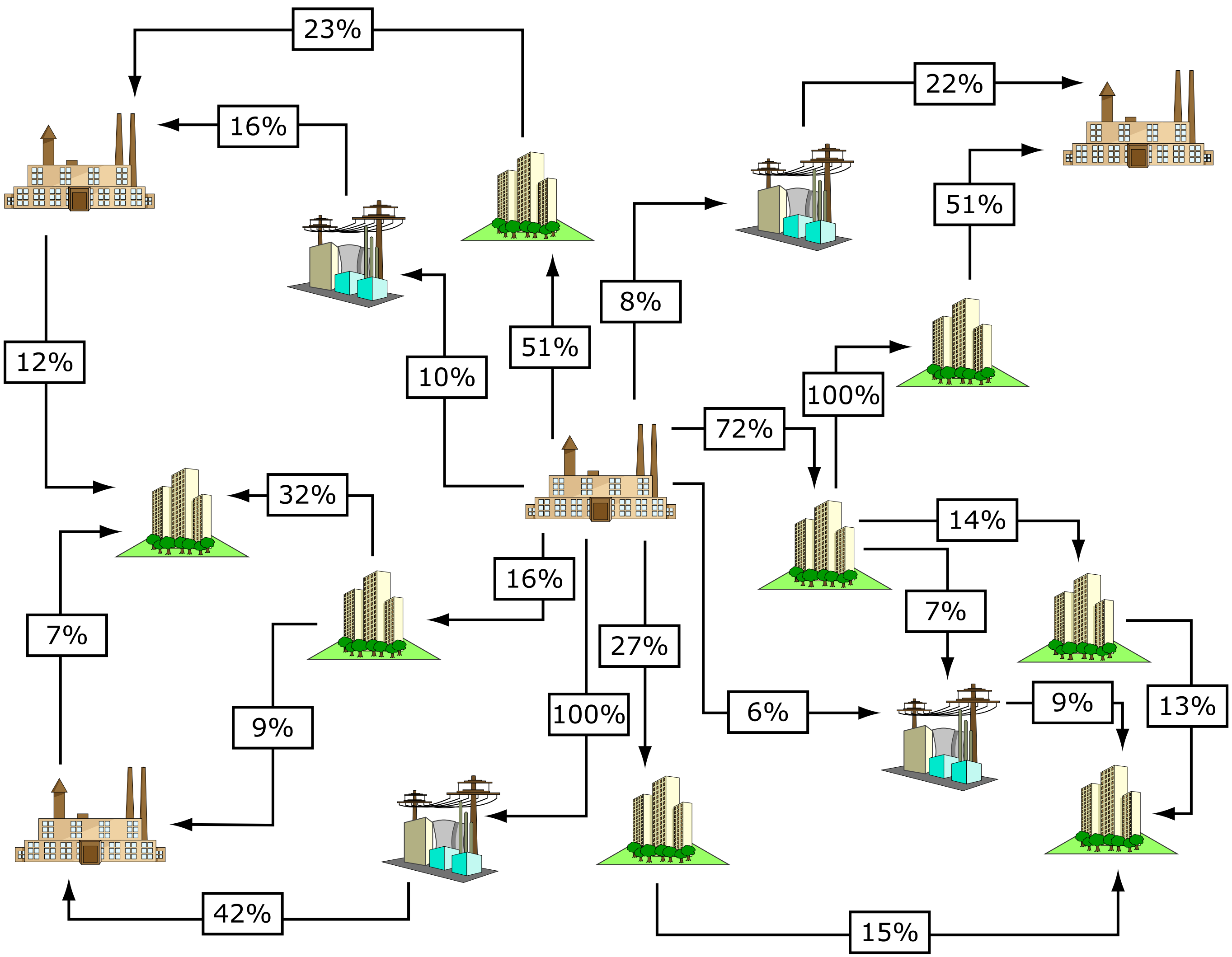
- Actors / nodes / vertices / points
- Ties / edges / arcs / lines / links
  - connect pair of actors
  - types of social relations
    - friendship
    - acquaintance
    - kinship
    - advice
    - hindrance
    - sex
  - allow different kind of flows
    - messages
    - money
    - diseases



What is a Social Network? - Relations among People



What is a Network? - Relations among Institutions



- as institutions
  - owned by, have partnership / joint venture
  - purchases from, sells to
  - competes with, supports
- through stakeholders
  - board interlocks
  - Previously worked for

Image by MIT OpenCourseWare.

Why study social networks?

Example 2) Homophily Theory

	Male	Female
Male	123	68
Female	95	164

- Birds of a feather flock together
- See McPherson, Smith-Lovin & Cook (2001)

	0-13	14-29	30-44	45-65	>65
0-13	212	63	117	72	91
14-29	83	372	75	67	84
30-44	105	98	321	214	117
45-65	62	72	232	412	148
>65	90	77	124	153	366

- age / gender → network

Managerial Relevance – Social Network...

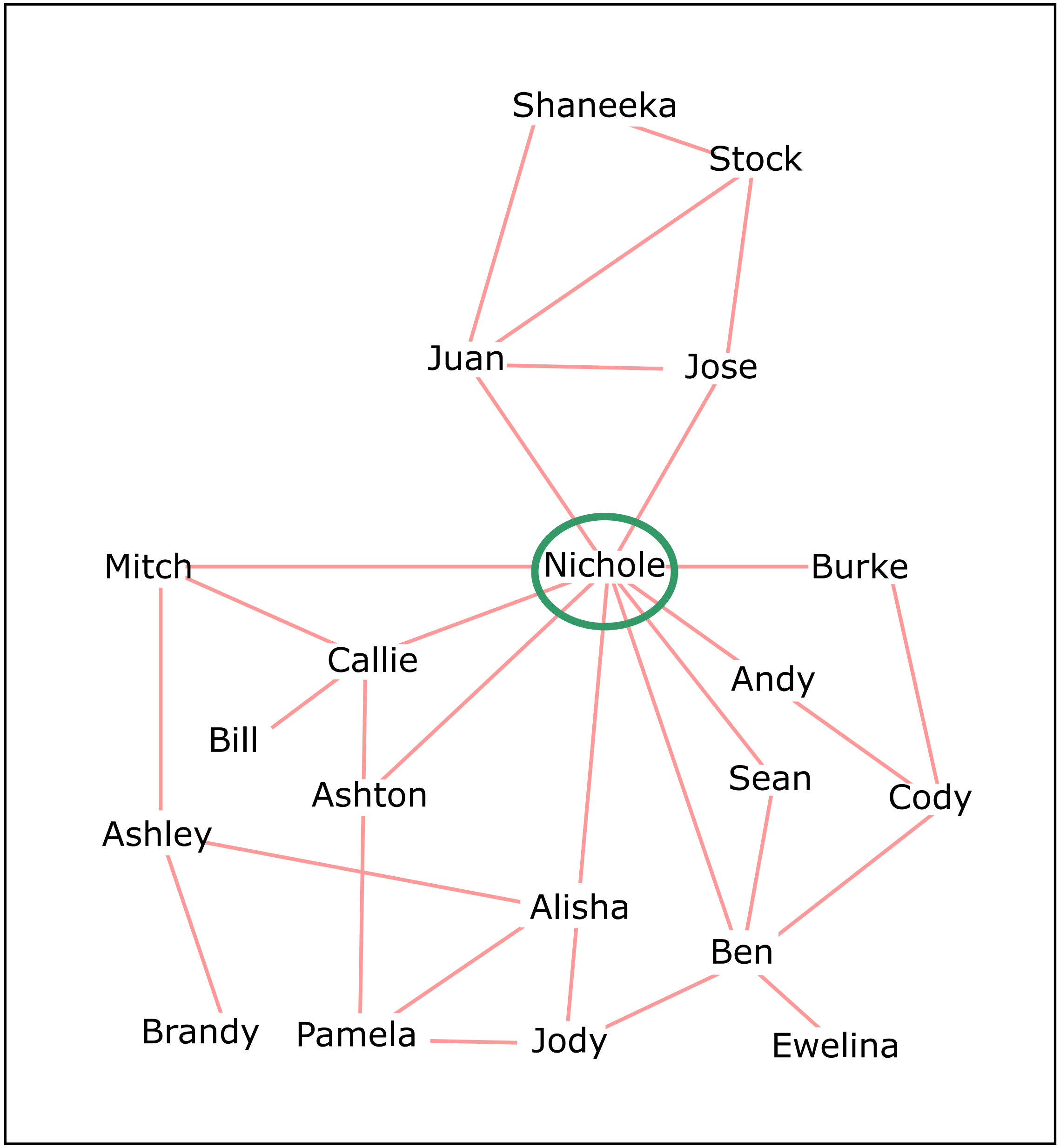


Image by MIT OpenCourseWare.

…vs. Organigram

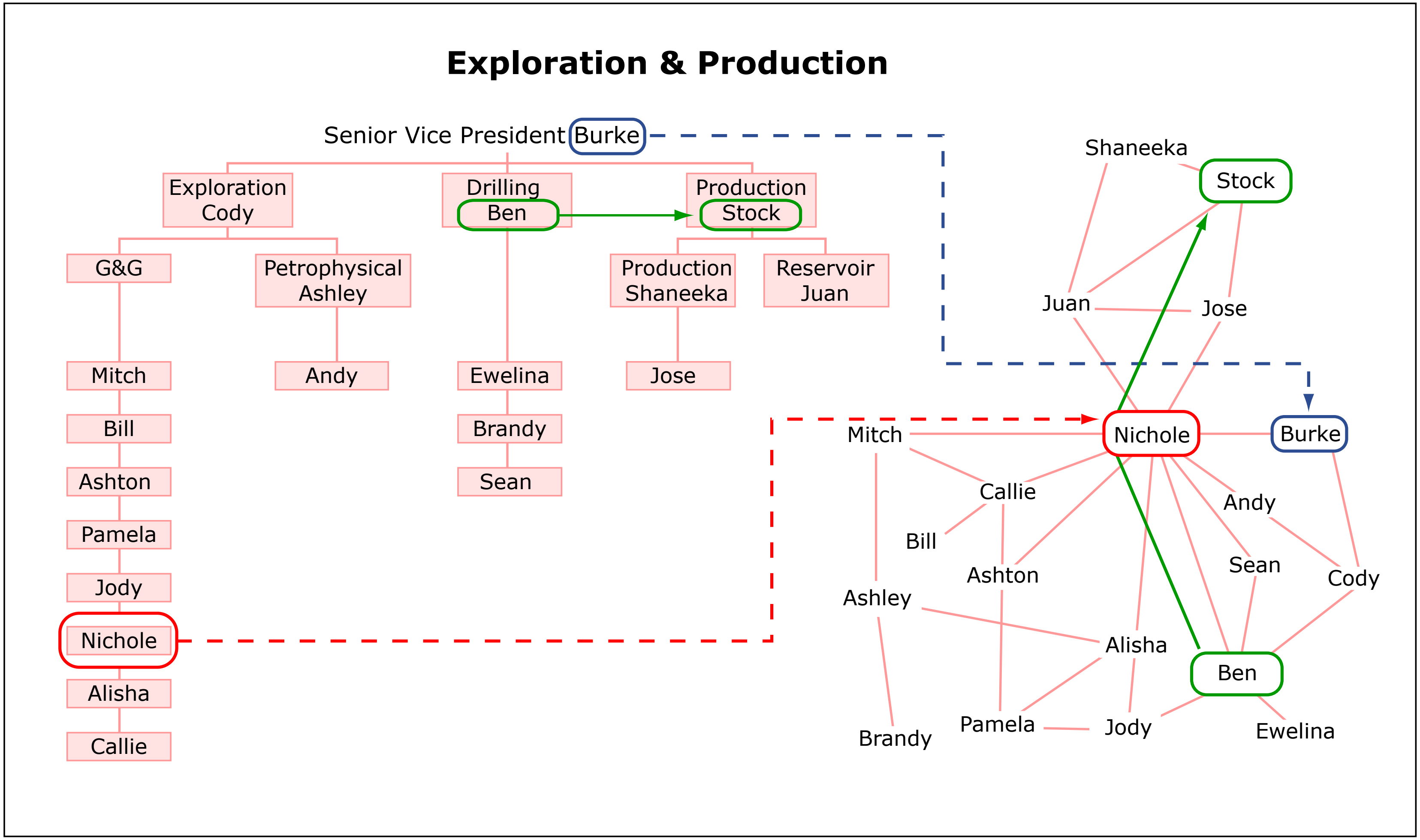
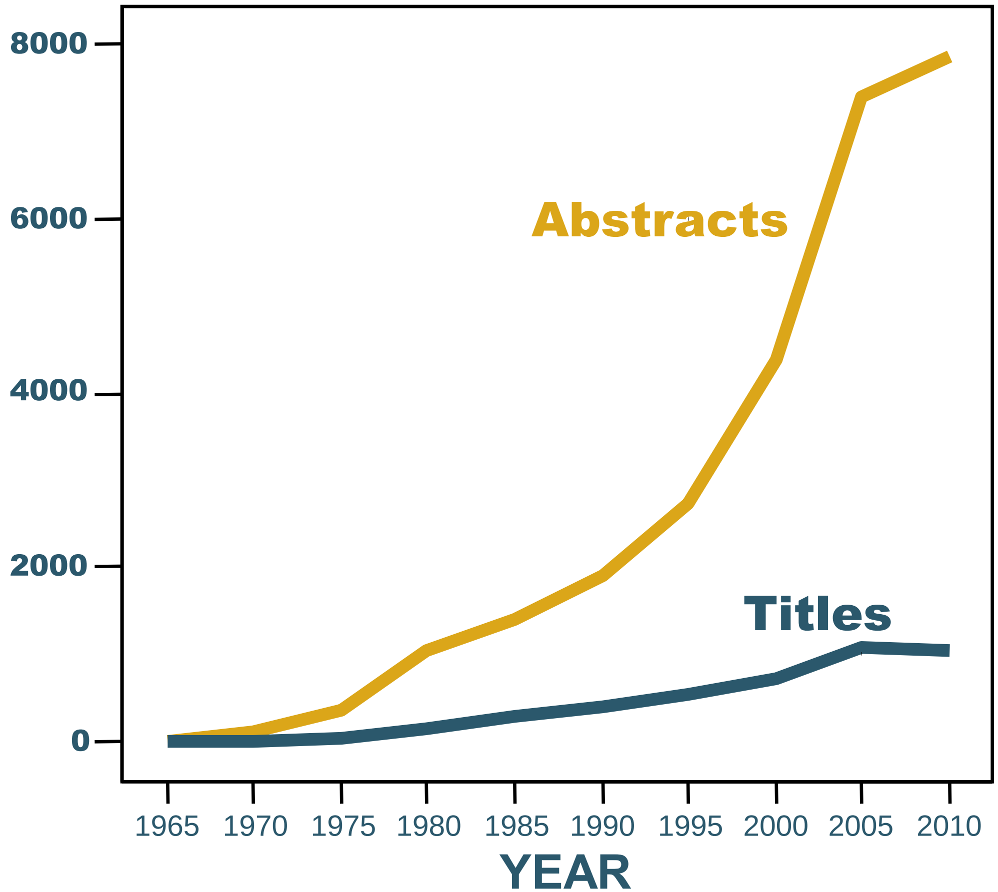


Image by MIT OpenCourseWare.

Source: http://www.robcross.org/sna.htm

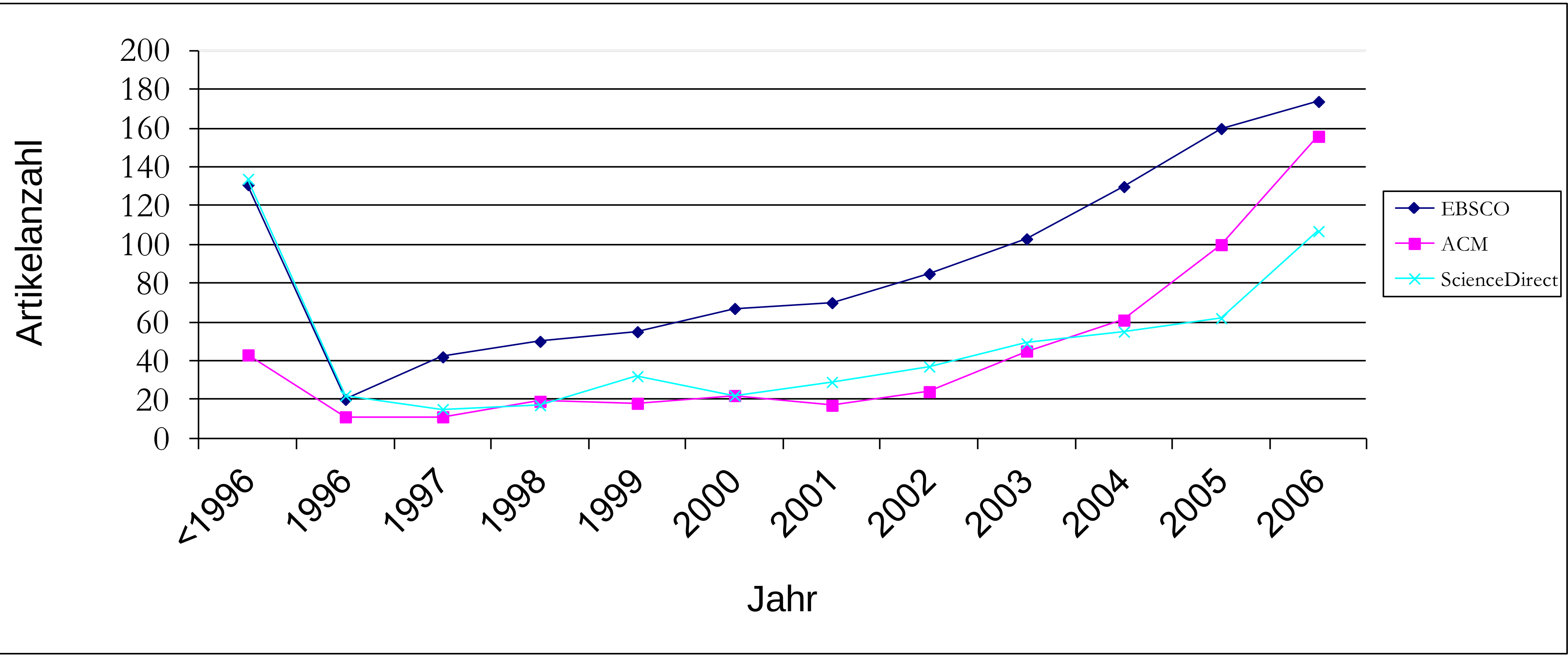
SNA – A Recent Trend in Social Sciences Research

- Keyword search for„social“ + „network“ in 14 literature databases



Source: *Knoke, David (2007) Introduction to Social Network Analysis*

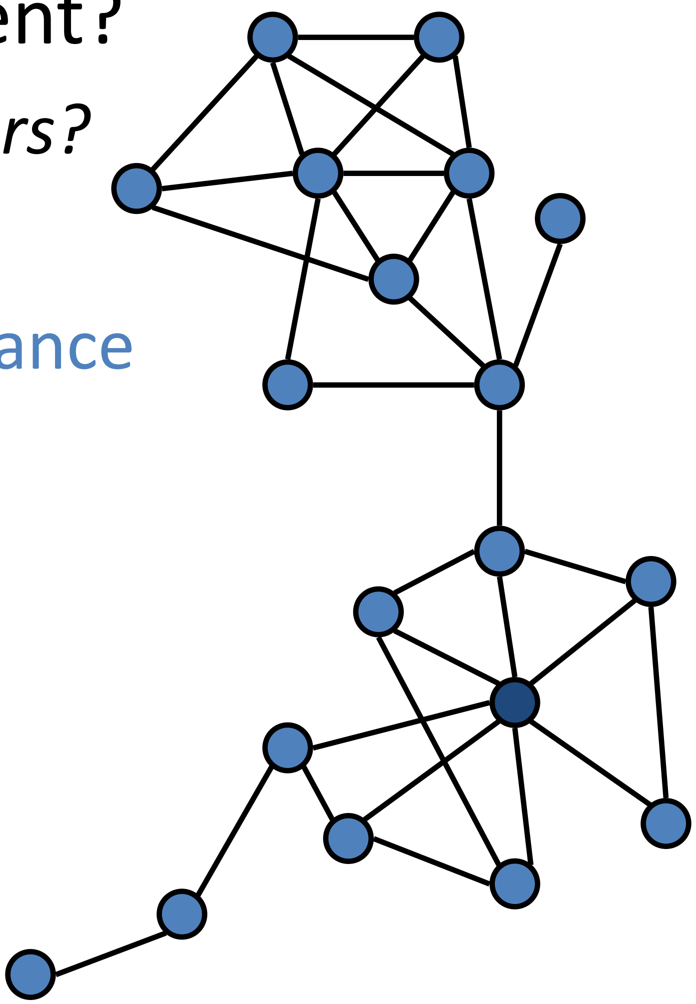
SNA – A Recent Trend in IS Research



How to analyze Social Networks?

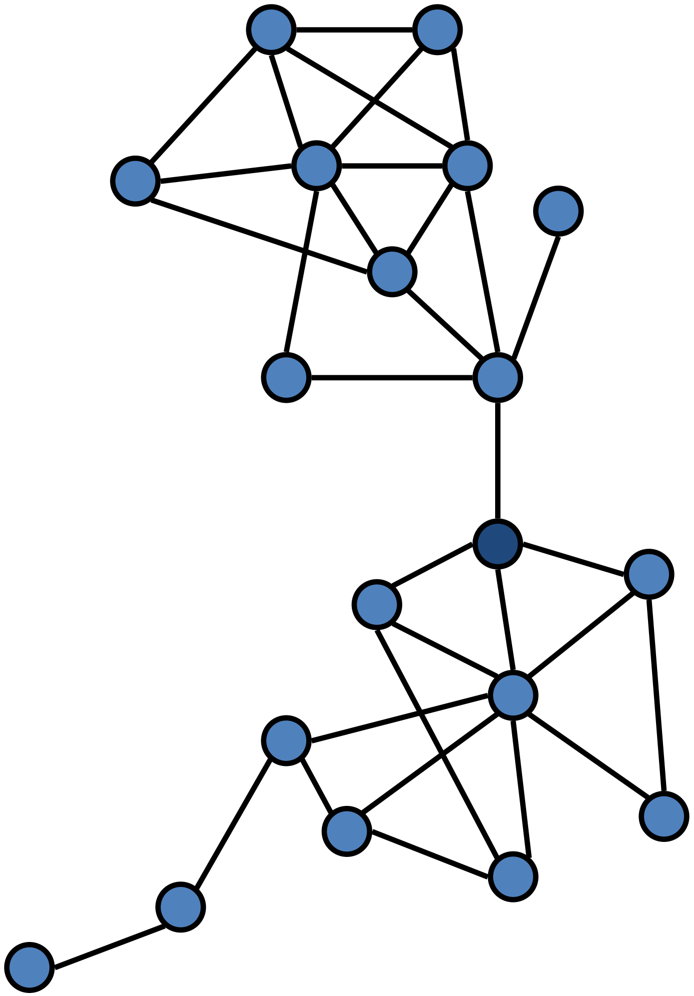
Example: Centrality Measures

- Who is the most prominent?
  - *Who knows the most actors? (Degree Centrality)*
  - Who has the shortest distance to the other actors?
  - Who controls knowledge flows?
  - ...



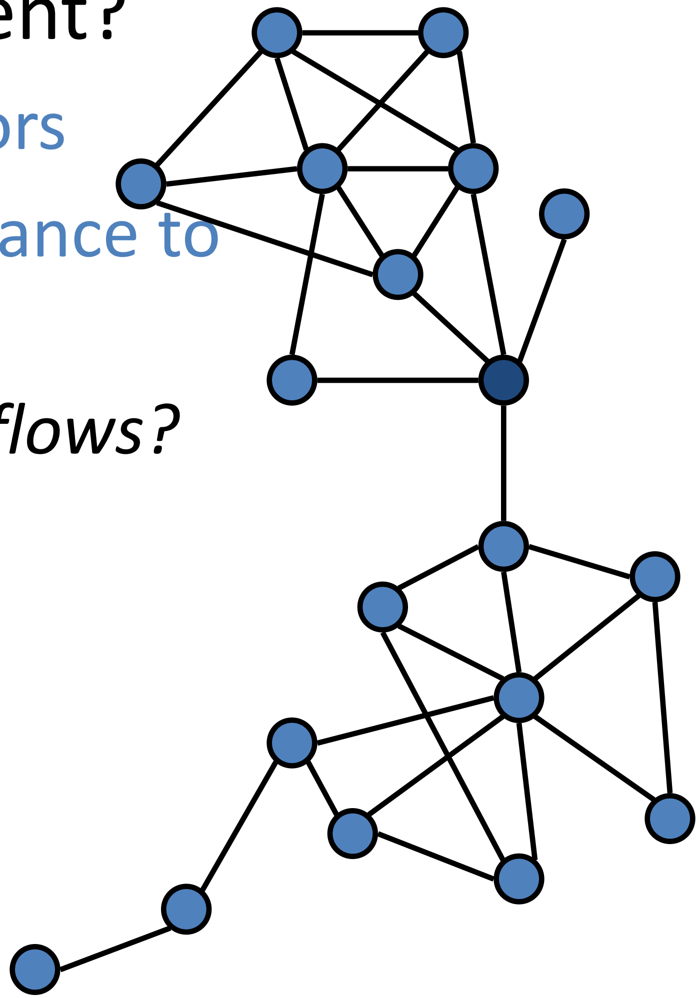
Example: Centrality Measures

- Who is the most prominent?
  - Who knows the most actors?
  - *Who has the shortest distance to the other actors? (Closesness Centrality)*
  - Who controls knowledge flows?
  - ...



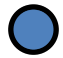
Example: Centrality Measures

- Who is the most prominent?
  - Who knows the most actors
  - Who has the shortest distance to the other actors?
  - *Who controls knowledge flows? (Betweenness Centrality)*
  - …



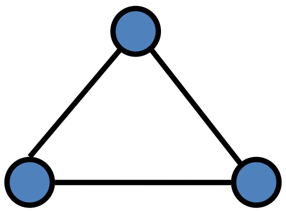
Basic Concepts

Dyads, Triads and Relations



- *actor*

- *dyad*



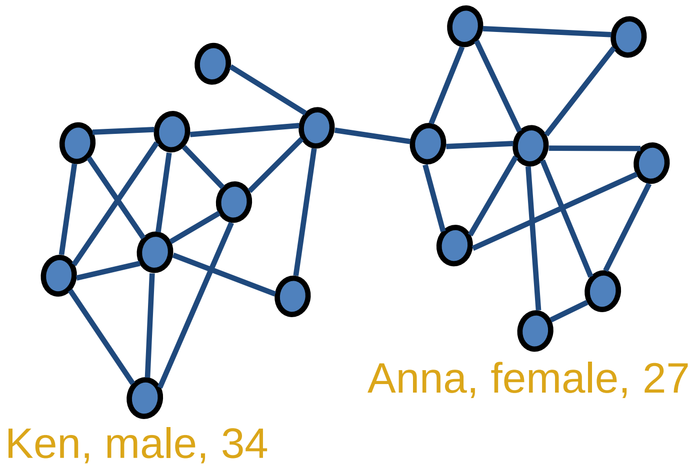
- *triad*

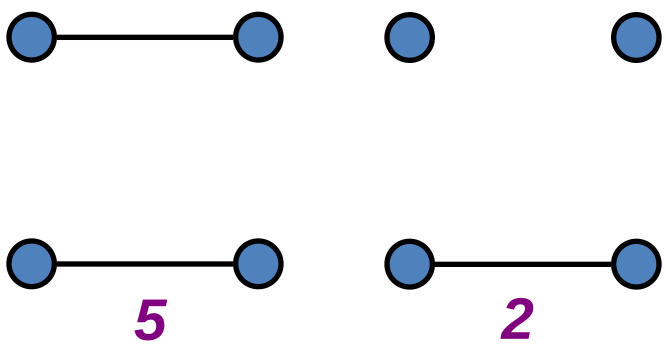
friendship

kinship

- *relation:*
  - collection of specific ties among members of a group

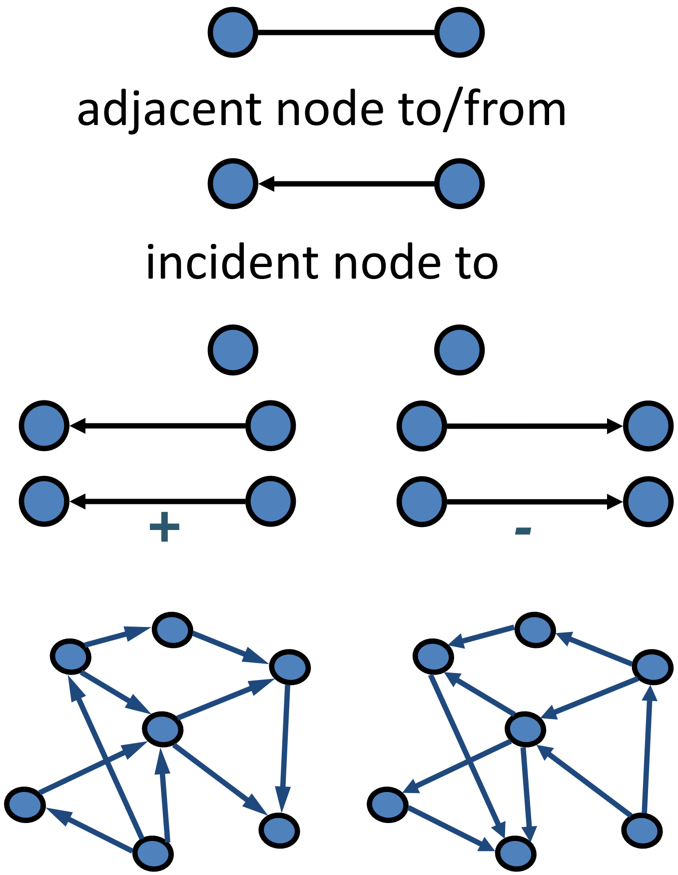
Strength of a Tie





- *Social network*
  - finite set of actors and relation(s) defined on them
  - depicted in *graph/ sociogram*
    - *labeled graph*
- *Strength of a Tie*
  - *dichotomous* vs. *valued*
    - depicted in *valued graph* or *signed graph* (+/-)

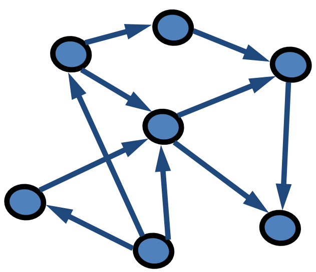
Strength of a Tie



- *Strength of a Tie*
  - *nondirectional* vs. *directional*
    - depicted in *directed graphs (digraphs)*
    - nodes connected by *arcs*
    - 3 *isomorphism classes*
      - *null dyad*
      - *mutual / reciprocal / symmetrical dyad*
      - *asymmetric / antisymmetric dyad*
    - *converse of a digraph*
      - *reverse direction of all arcs*

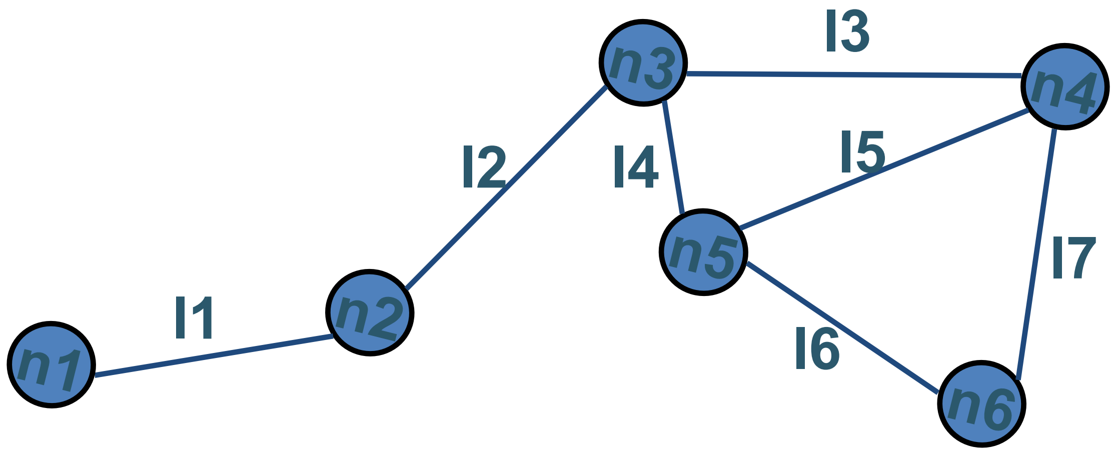
Walks, Trails, Paths

- *(Directed) Walk (W)*
  - sequence of nodes and lines starting and ending with (different) nodes (called *origin* and *terminus*)
  - Nodes and lines can be included more than once
- *Inverse* of a (directed) walk *(W^{-1})*
  - Walk in opposite order



- *Length* of a walk
  - How many lines occur in the walk? (same line counts double, in weighted graphs add line weights)
- *(Directed) Trail*
  - Is a walk in which all lines are distinct
- *(Directed) Path*
  - Walk in which all nodes and all lines are distinct
- Every path is a trail and every trail is a walk

Walks, Trails and Paths - Repetition



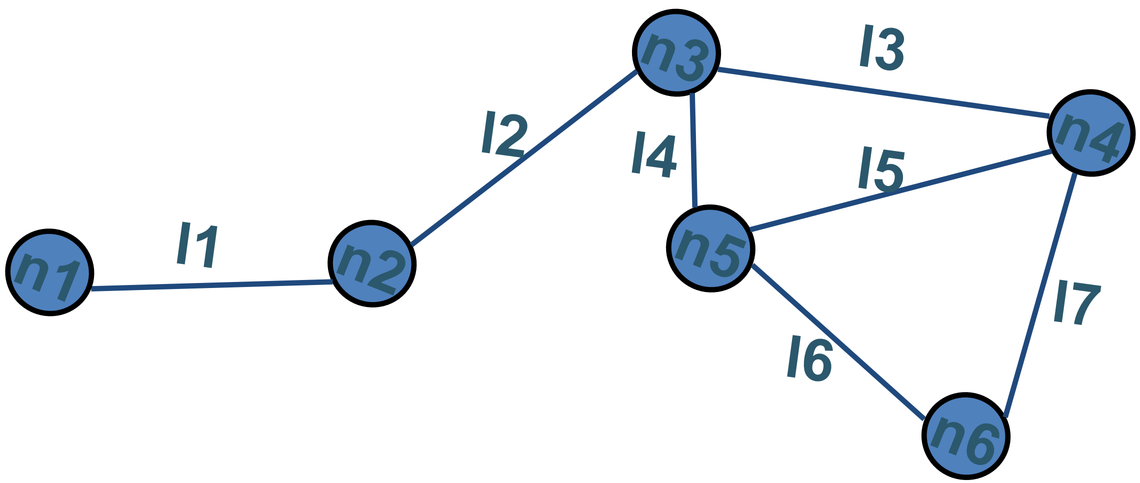
- W = n1 l1 n2 l2 n3 l4 n5 l6 n6
  - n1
  - n3
- W = n1 l1 n2 l2 n3 l4 n5 l4 n3
- W = n1 l1 n2 l2 n3 l4 n5 l5 n4 l3 n3

- Path
  - origin
  - terminus
- Walk
- Trail

University of Cologne
Johannes Putzke

Reachability, Distances and Diameter

- *Reachability*
  - If there is a path between nodes n_i and n_j
- *Geodesic*
  - Shortest path between two nodes
- *(Geodesic) Distance* d(i,j)
  - Length of Geodesic (also called „degrees of separation")



Mathematical Notation and Fundamentals

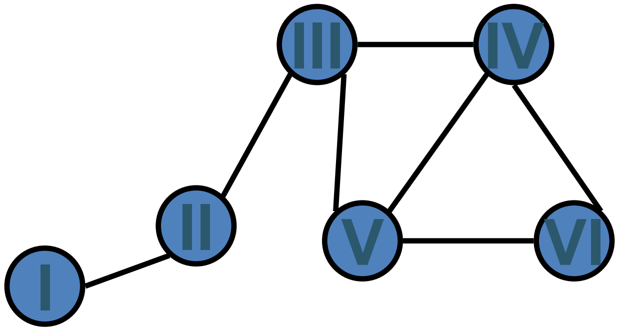
Three different notational schemes

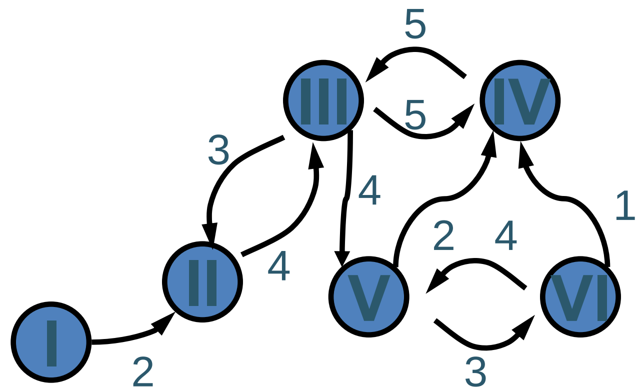
1. Graph theoretic
2. Sociometric
3. Algebraic

1. Graph Theoretic Notation

- N Actors $\{n_1, n_2, \ldots, n_g\}$
- $n_1 \rightarrow n_j$ there is a tie between the ordered pair $\langle n_i, n_j \rangle$
- $n_1 \rightarrow n_j$ there is no tie
- (n_i, n_j) nondirectional relation
- $\langle n_i, n_j \rangle$ directional relation
- $g(g-1)$ number of ordered pairs in $\langle n_i, n_j \rangle$ directional network
- $g(g-1)/2$ number of ordered pairs in nondirectional network
- L collection of ordered pairs with ties $\{l_1, l_2, \ldots, l_g\}$
- G graph descriped by sets (N, L)
- *Simple graph* has no *reflexive ties, loops*

2. Sociometric Notation - From Graphs to (Adjacency/Socio)-Matrices





Binary, undirected

	I	II	III	IV	V	VI
I	-					
II	1	-		symmetrical		
III		1	-			
IV			1	-		
V			1	1	-	
VI				1	1	-

Valued, directed

	I	II	III	IV	V	VI
I		2	0	0	0	0
II	0	0	4	0	0	0
III	0	3	0	5	4	0
IV	0	0	5	0	0	0
V	0	0	0	2	0	3
VI	0	0	0	1	4	0

2. Sociometric Notation

- X g × g *sociomatrix* on a single relation
  g × g × R *super-sociomatrix* on R relations
    - X_R sociomatrix on relation R
- $X_{ij(r)}$ value of tie from n_i to n_i (on relation χ_r) where i ≠ j

2. Sociometric Notation – From Matrices to Adjacency Lists and Arc Lists

	I	II	III	IV	V	VI
I	-	1				
II	1	-	1			
III		1	-	1	1	
IV			1	-	1	1
V			1	1	-	1
VI				1	1	-

Adjacency List

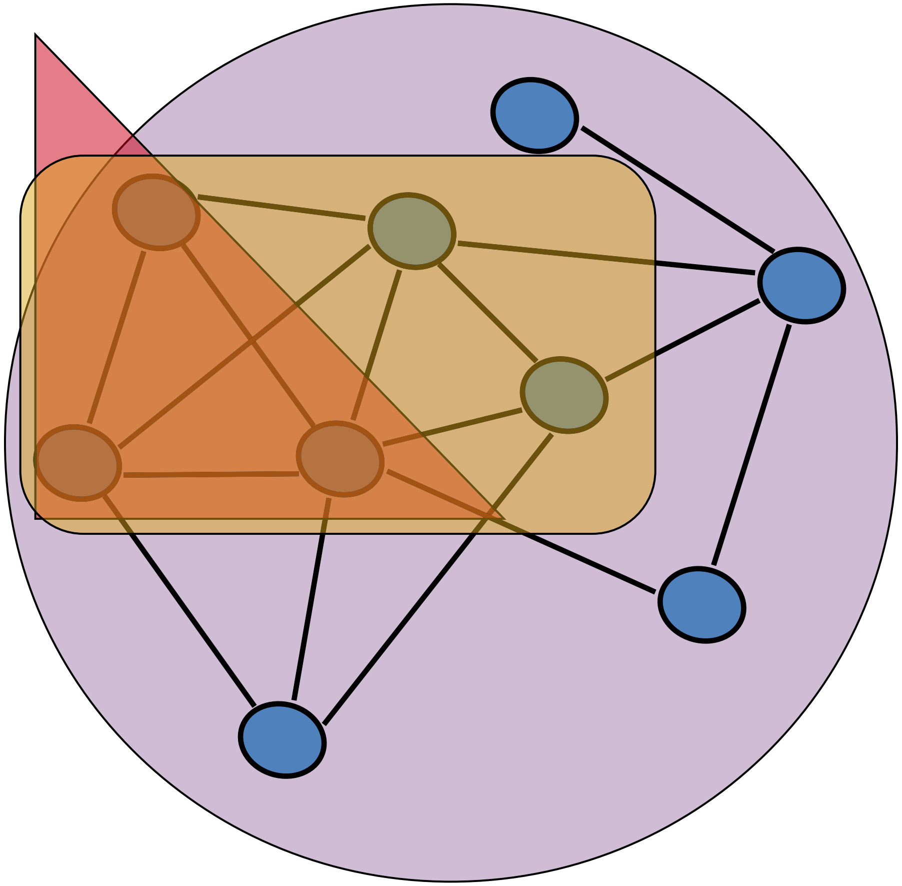
I	II
II	I III
III	II IV V
IV	III V VI
V	III V VI
VI	IV V

Arc List

I II
II I
II III
III II
III IV
III V
IV III
IV V
IV VI
V III
V IV
V VI
VI IV
VI IV

Network Statistics

Different Levels of Analysis



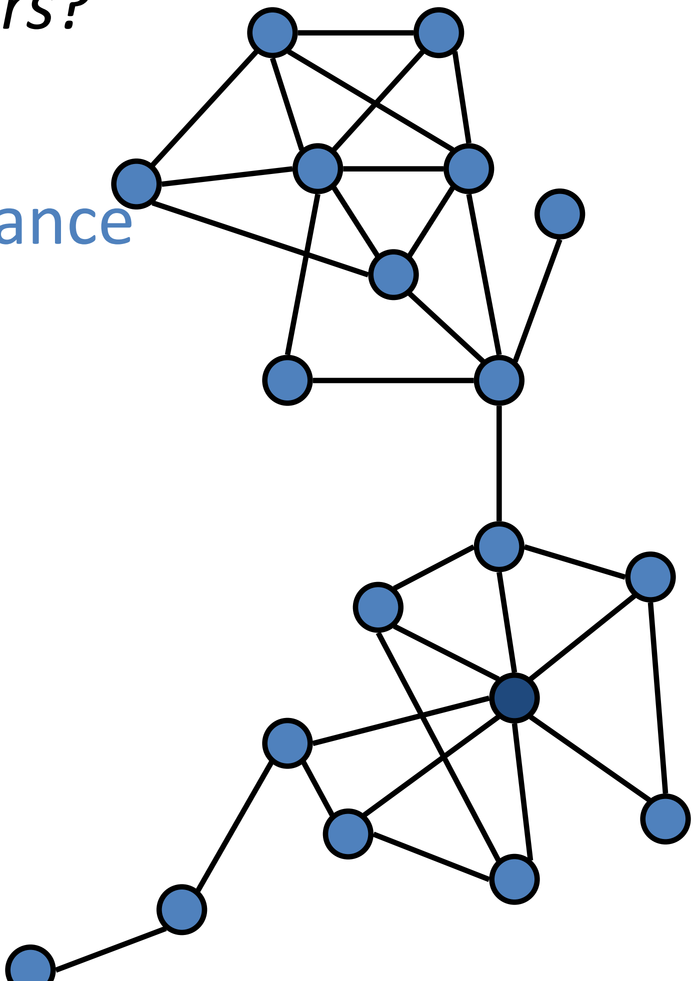
- Actor-Level
- Dyad-Level
- Triad-Level
- Subset-level (cliques / subgraphs)
- Group (i.e. global) level

Measures at the Actor-Level:

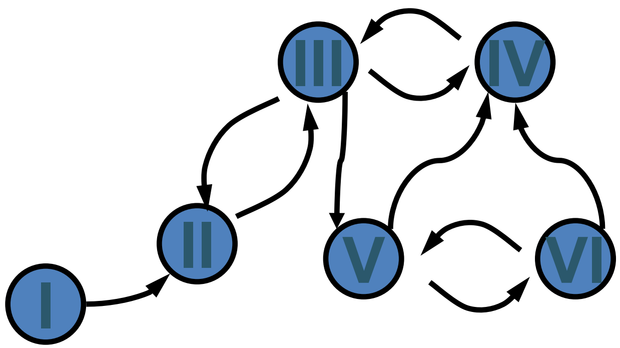
Measures of Prominence: Centrality and Prestige

Degree Centrality

- *Who knows the most actors? (Degree Centrality)*
- Who has the shortest distance to the other actors?
- Who controls knowledge flows?
- ...



Degree Centrality I



	I	II	III	IV	V	VI	
I		1	0	0	0	0	1
II	0	0	1	0	0	0	1
III	0	1	0	1	1	0	3
IV	0	0	1	0	0	0	1
V	0	0	0	1	0	1	2
VI	0	0	0	1	0	1	2
	0	2	2	3	1	2	

- *Indegree $d_I(n_i)$*
  - *Popularity, status, deference, degree prestige*

$$C_{DI}\left(n_i\right) = d_I\left(n_i\right) = \sum_j^g x_{ji} = x_{+i}$$

- *Outdegree $d_O(n_i)$*
  - Expansiveness

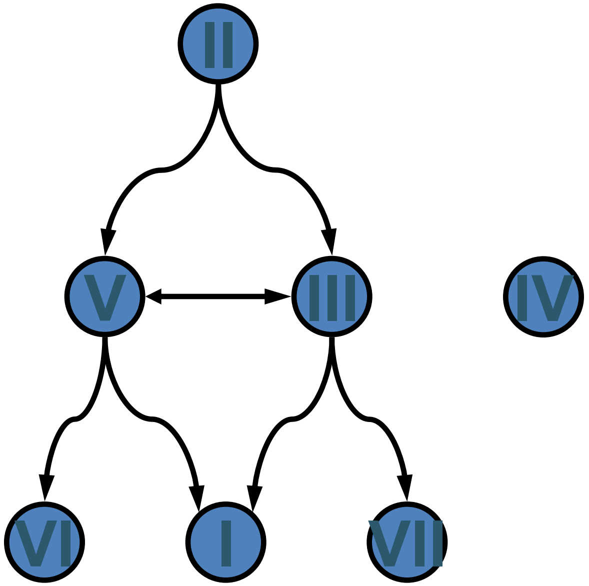
$$C_{DO}\left(n_i\right) = d_o\left(n_i\right) = \sum_j^g x_{ij} = x_{i+}$$

- Total degree ≡ 2 x number of edges

Marginals of adjacency matrix

Degree Centrality II

- Interpretation: opportunity to (be) influence(d)
- Classification of Nodes
  - *Isolates*
    - $d_I(n_i) = d_O(n_i) = 0$
  - *Transmitters*
    - $d_I(n_i) = 0$ and $d_O(n_i) > 0$
  - *Receivers*
    - $d_I(n_i) > 0$ and $d_O(n_i) = 0$
  - *Carriers / Ordinaries*
    - $d_I(n_i) > 0$ and $d_O(n_i) > 0$

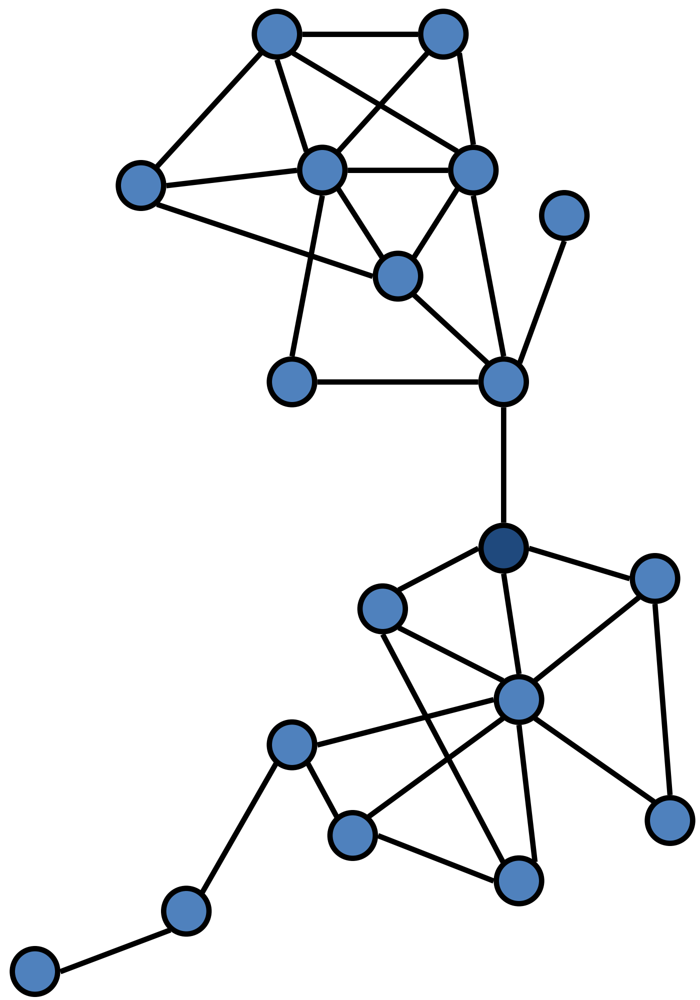


- Standardization of C_D to allow comparison across networks of different sizes: divide by ist maximum value

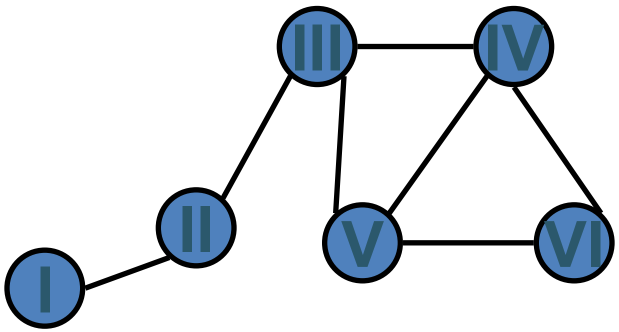
$$C_D^{'}(n_i) = \frac{d(n_i)}{g-1}$$

Closeness Centrality

- Who knows the most actors?
- *Who has the shortest distance to the other actors? (Closesness Centrality)*
- Who controls knowledge flows?
- ...



Closeness Centrality



	I	II	III	IV	V	VI	
I	-	1	2	3	3	4	13
II	1	-	1	2	2	3	9
III	2	1	-	1	1	2	7
IV	3	2	1	-	1	1	8
V	3	2	1	1	-	1	8
VI	4	3	2	1	1	-	11

- Index of expected arrival time

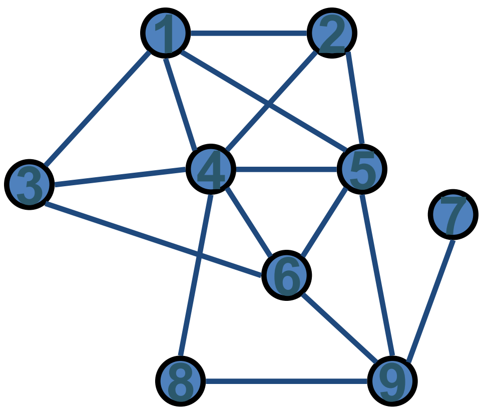
$$C_C(n_i) = \frac{1}{\sum_{j=1}^{g} d\left(n_i, n_j\right)}$$

Reciprocal of marginals of geodesic distance matrix

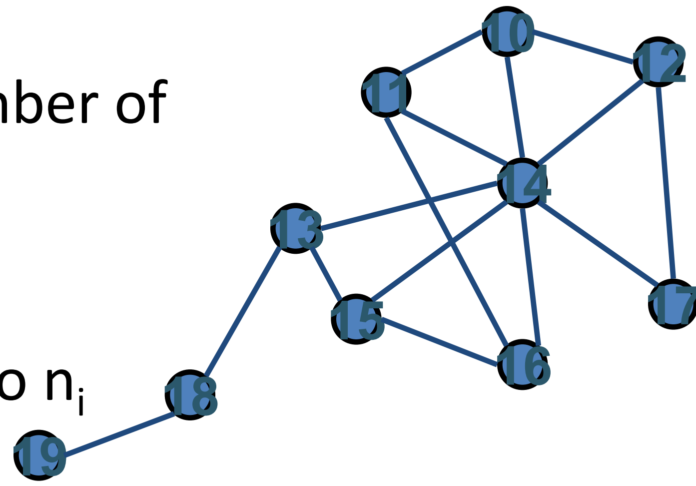
- Standardize by multiplying (g-1)
- Problem: Only defined for connected graphs

Proximity Prestige

$$P_P(n_i) = \frac{I_i / (g-1)}{\sum_{j=1}^{g} d\left(n_j, n_i\right) / I_i}$$

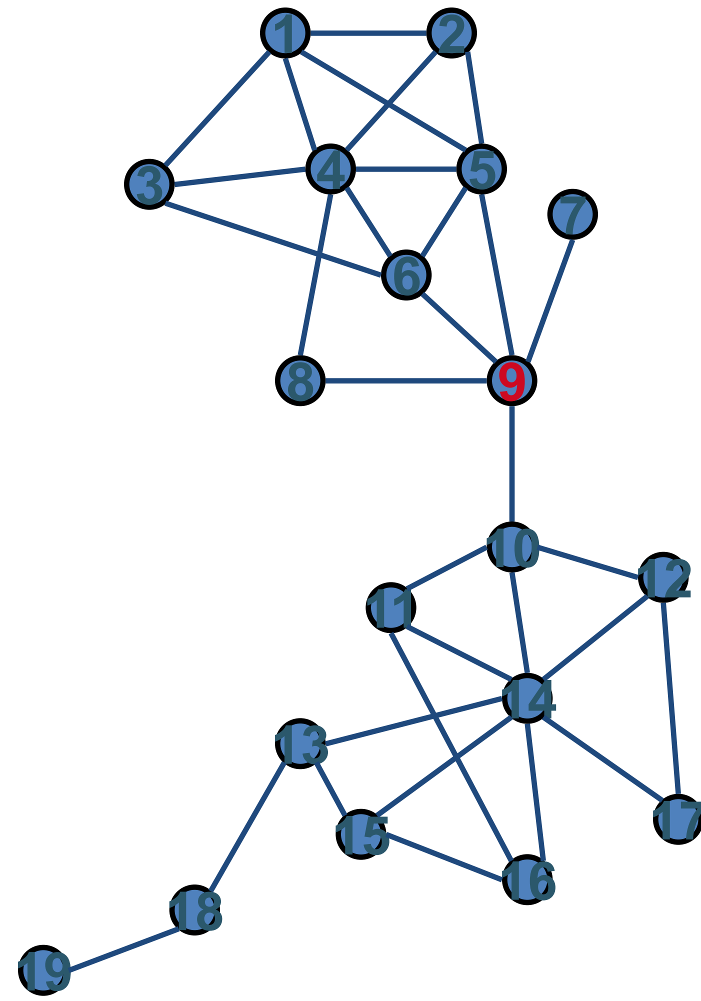


- $I_i/(g-1)$
  - number of actors in the influence domain of n_i
  - normed by maximum possible number of actors in influence domain
- $\Sigma d(n_j,n_i)/ I_i$
  - average distance these actors are to n_i



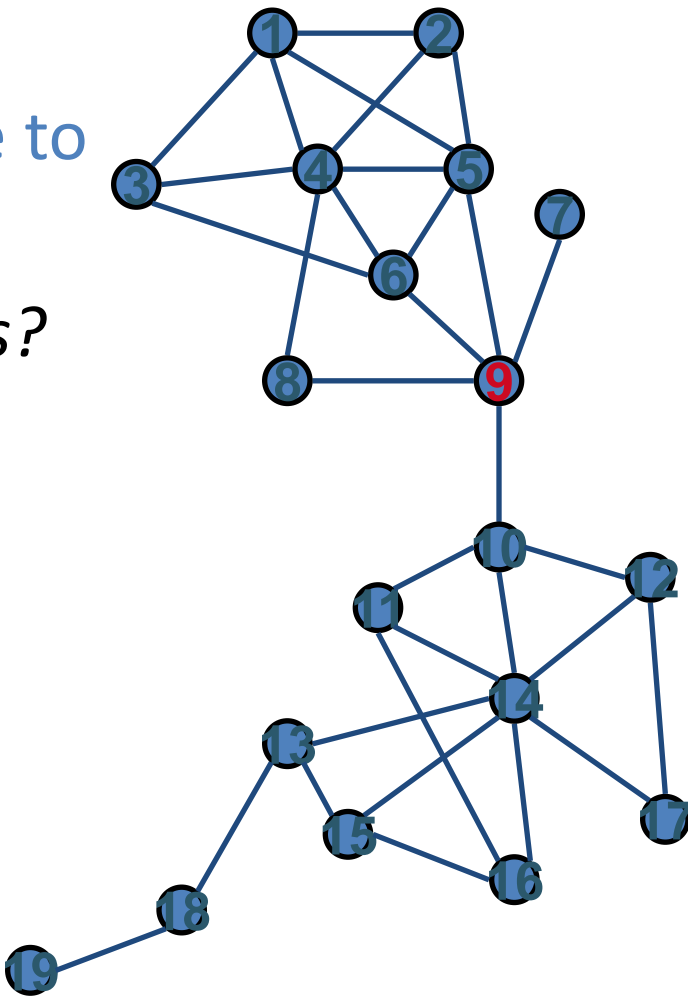
Eccentricity / Association Number

- Largest geodesic distance between a node and any other node
- $\max_j d(i,j)$



Betweenness Centrality

- Who knows the most actors?
- Who has the shortest distance to the other actors?
- *Who controls knowledge flows? (Betweenness Centrality)*
- ...

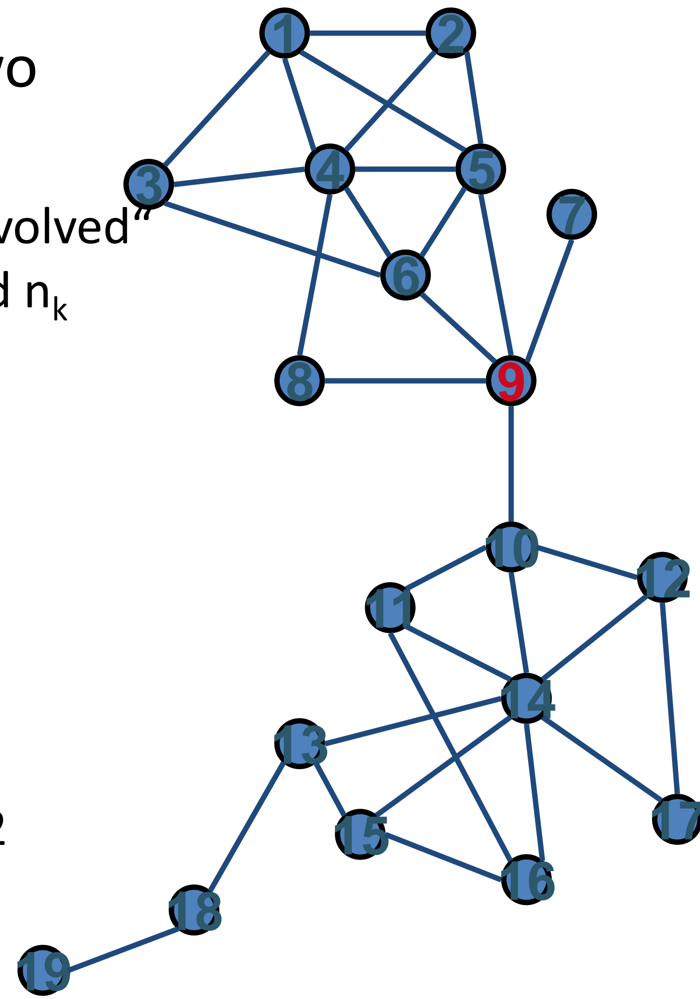


Betweenness Centrality

- How many geodesic linkings between two actors j and k contain actor i?
  - $g_{jk}(n_i)/g_{jk}$ probability that distinct actor n_i „involved“ in communication between two actors n_j and n_k

$$C_B(n_i) = \frac{\sum_{j<k} g_{jk}(n_i)}{g_{jk}}$$

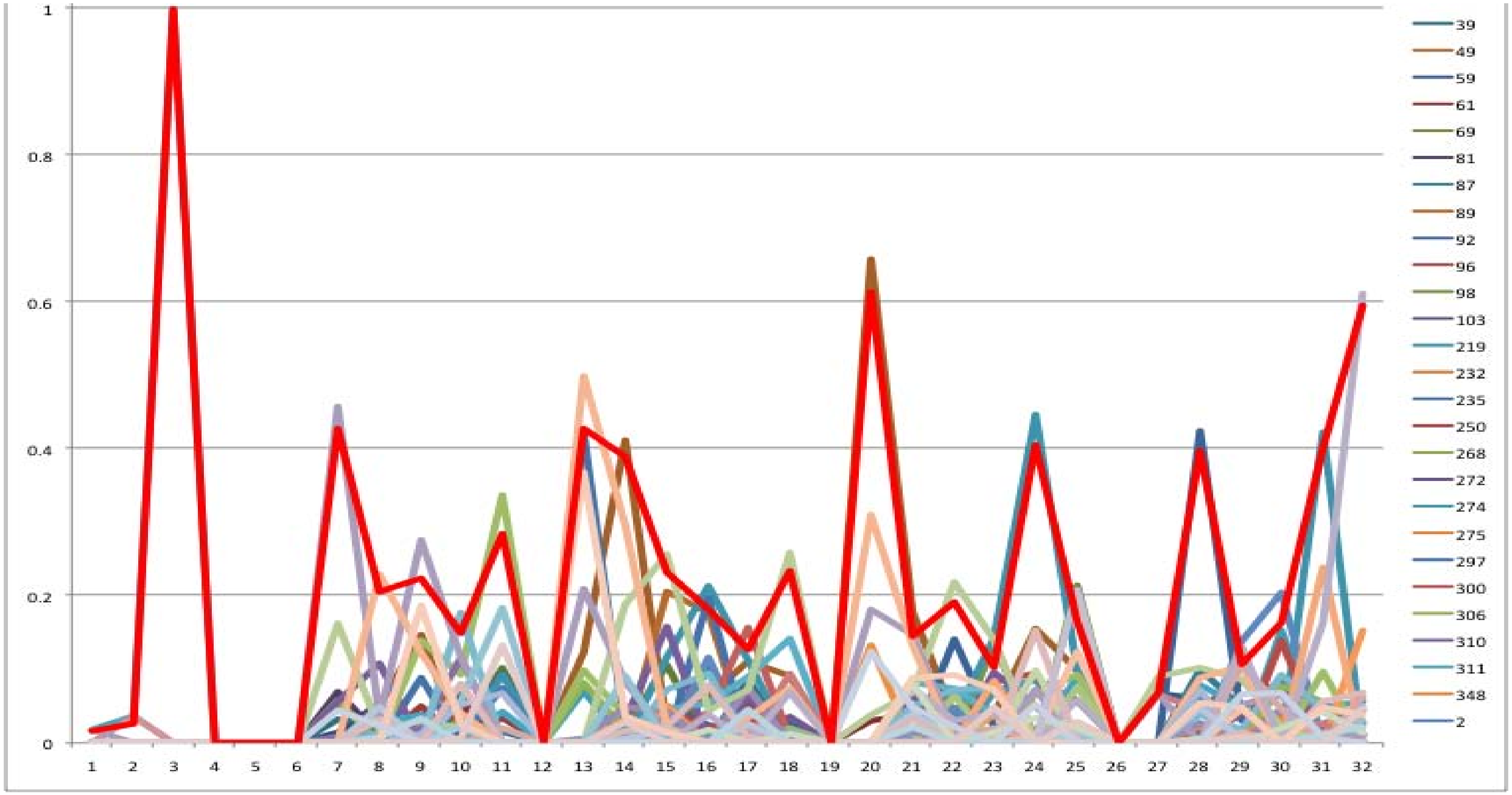
  - standardized by dividing through (g-1)(g-2)/2



Several other Centrality Measures

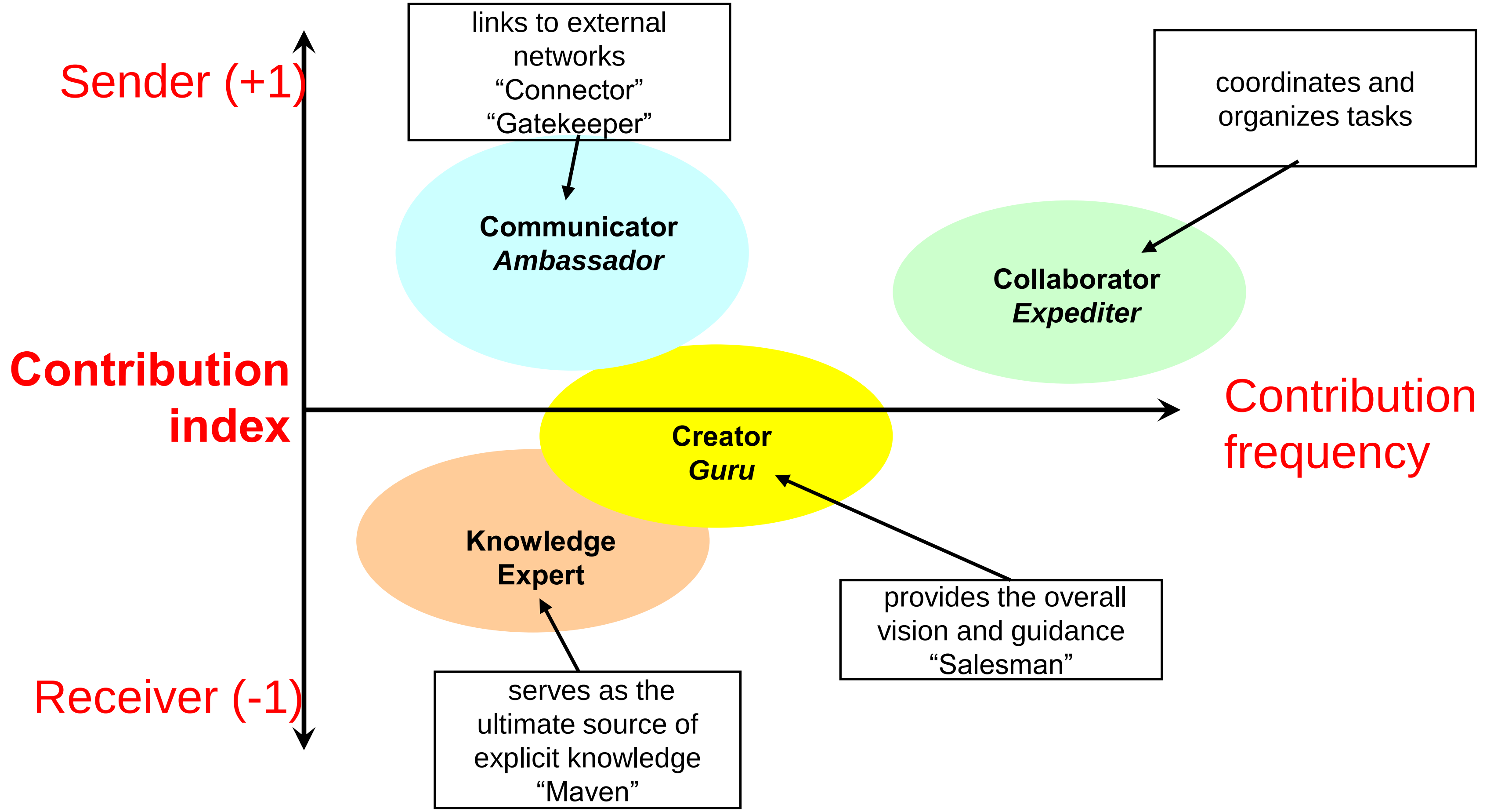
- …beyond the scope of this lecture
  - *Status or Rank Prestige, Eigenvector Centrality*
    - also reflects status or prestige of people whom actor is linked to
    - Appropriate to identify *hubs* (actors adjacent to many peripheral nodes) and *bridges* (actors adjacent to few central actors)
      - attention: more common, different meaning of bridge!!!
  - *Information Centrality*
  - see Wasserman & Faust (1994), p. 192 ff.
  - *Random Walk Centrality*
    - see Newman (2005)

Condor – Betweenness Centrality



(Actor) Contribution Index

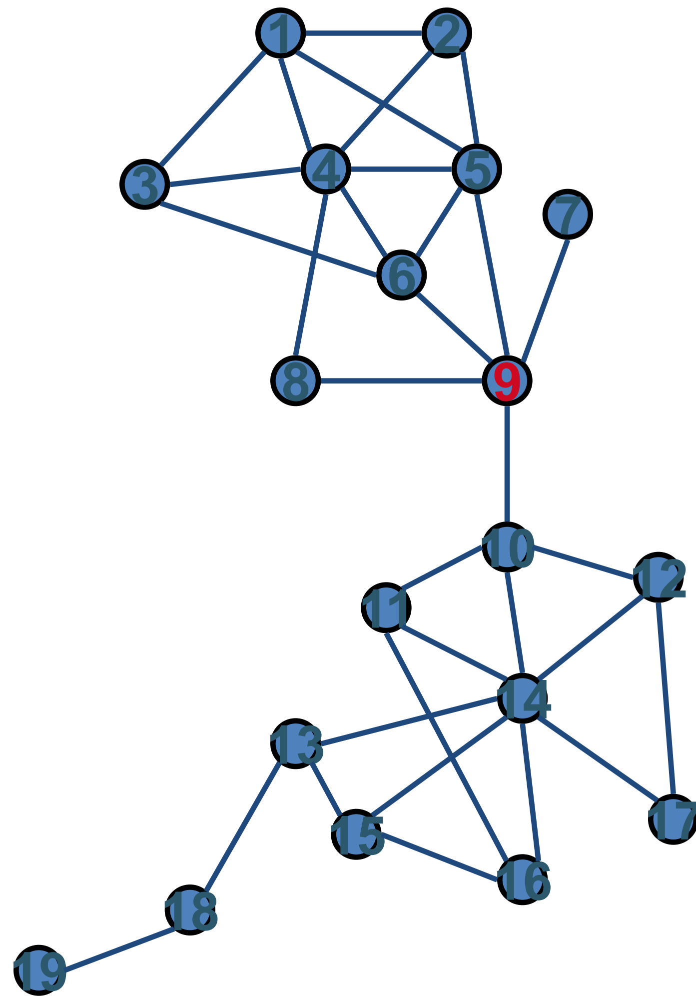
$$\frac{messages\ sent - messages\ received}{messages\ sent + messages\ received}$$



Measures at the Group-(Global-)Level and Subgroup-Level

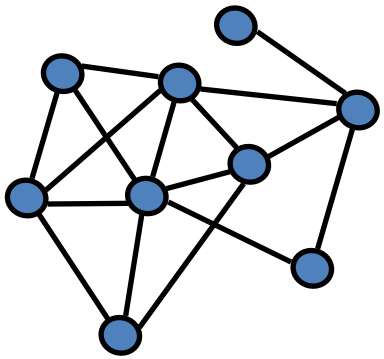
Diameter of a Graph and Average Geodesic Distance

- Diameter
  - Largest geodesic distance between any pair of nodes
- Average Geodesic Distance
  - How fast can information get transmitted?

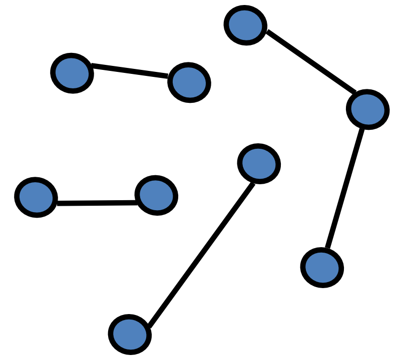


Density

- Proportion of ties in a graph

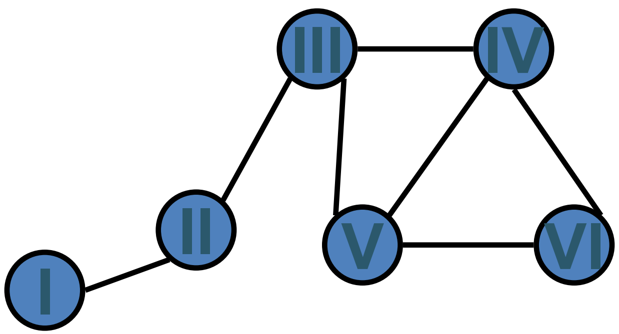


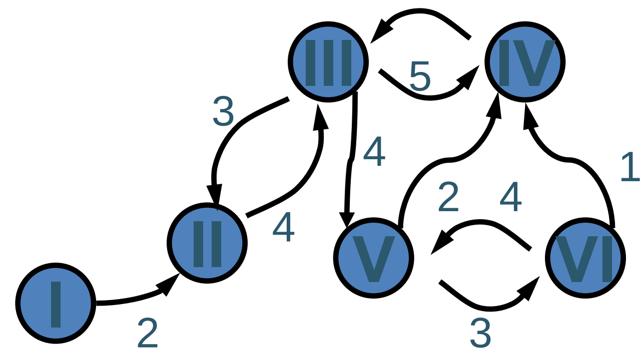
High density (44%)



Low density (14%)

Density





$$\Delta = \frac{L}{g(g-1)/2} = \frac{L}{\binom{g}{2}}$$

$$\Delta = \frac{\sum_{i=1}^{g}\sum_{j=1}^{g} x_{ij}}{g(g-1)}$$

In undirected graph:

Proportion of ties

In valued directed graph:

Average strength of the arcs

Group Centralization I

- How equal are the individual actors' centrality values?
  - $C_A(n_i^*)$ actor centrality index
  - $C_A(n^*)$ $\max_i C_A(n_i^*)$
  - $\sum_{i=1}^{g}\left[C_A(n^*)-C_A(n_i)\right]$ sum of difference between largest value and observed values
- General centralization index:

$$C_A = \frac{\sum_{i=1}^{g}\left[C_A(n^*)-C_A(n_i)\right]}{\max \sum_{i=1}^{g}\left[C_A(n^*)-C_A(n_i)\right]}$$

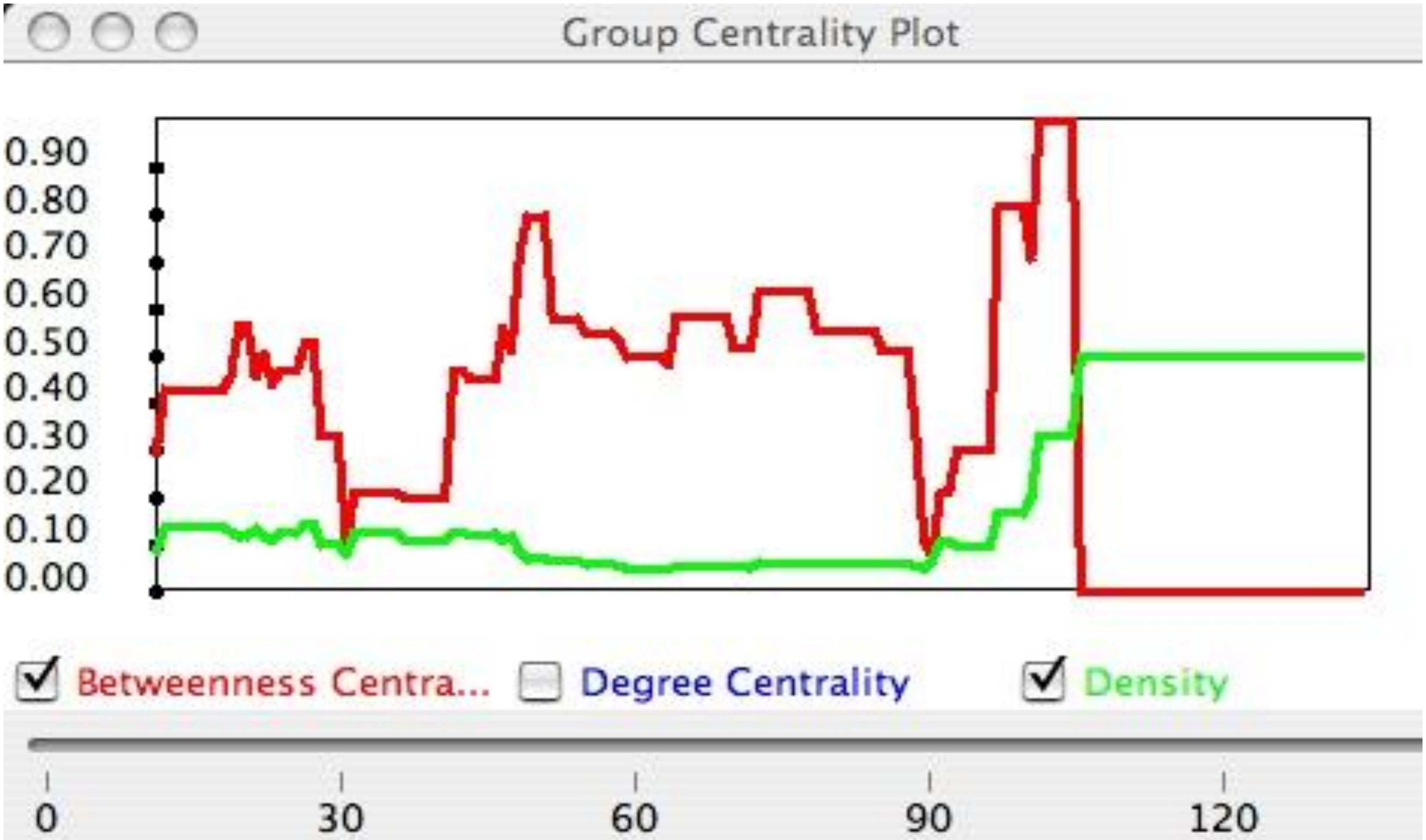
Group Centralization II

$$C_D = \frac{\sum_{i=1}^{g}\left[C_D\left(n^*\right) - C_D\left(n_i\right)\right]}{(g-1)(g-2)}$$

$$C_C = \frac{\sum_{i=1}^{g}\left[C_C^{'}\left(n^*\right) - C_C^{'}\left(n_i\right)\right]}{\left[(g-1)(g-2)\right](2g-3)}$$

$$CB = \frac{\sum_{i=1}^{g}\left[C_B\left(n^*\right) - C_B\left(n_i\right)\right]}{(g-1)^2(g-2)} = \frac{\sum_{i=1}^{g}\left[C_B^{'}\left(n^*\right) - C_B^{'}\left(n_i\right)\right]}{(g-1)}$$

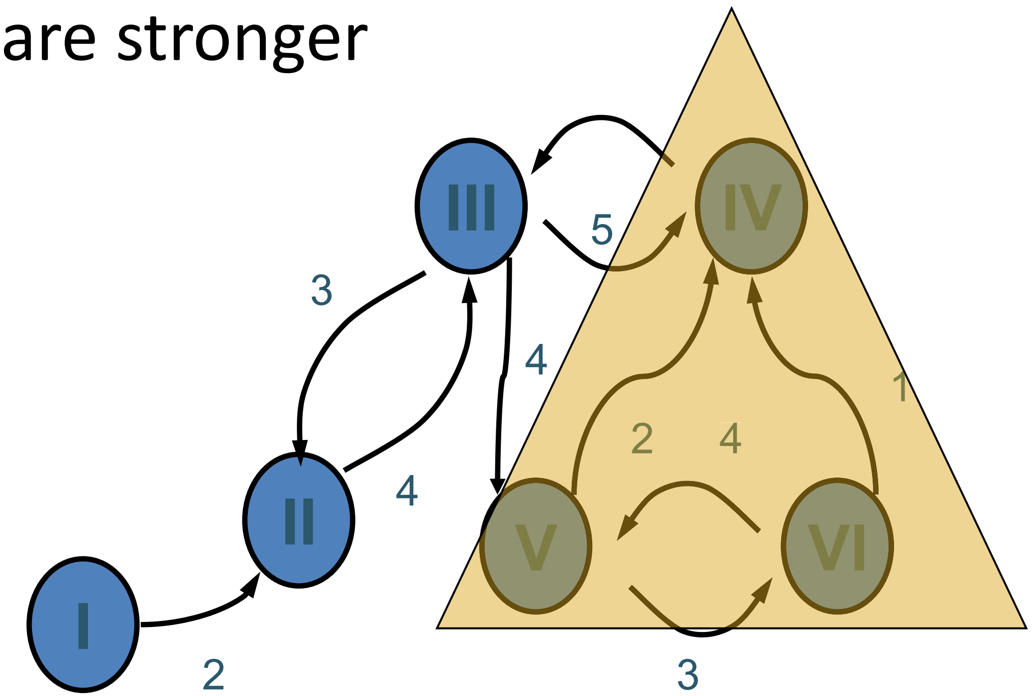
Condor – Group Centralization



Subgroup Cohesion

- average strength of ties within the subgroup divided by average strength of ties that are from subgroup members to outsiders
- >1 → ties in subgroup are stronger

$$\frac{\dfrac{\sum_{i \in N_s} \sum_{j \in N_s} x_{ij}}{g_s(g_s - 1)}}{\dfrac{\sum_{i \in N_s} \sum_{j \notin N_s} x_{ij}}{g_s(g - g_s)}}$$

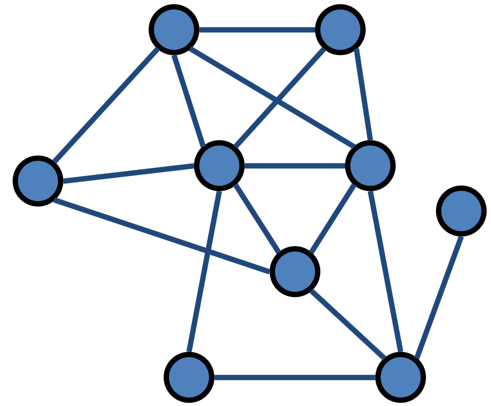


Connectivity of Graphs and Cohesive Subgroups

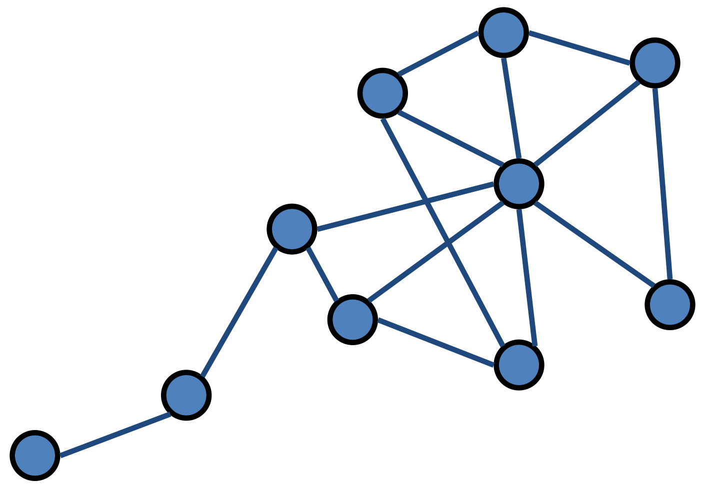
Connectivity of Graphs

Connected Graphs, Components, Cutpoints and Bridges

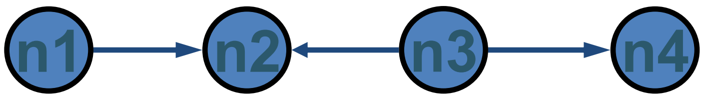
- *Connectedness*
  - A graph is connected if there is a path between every pair of nodes

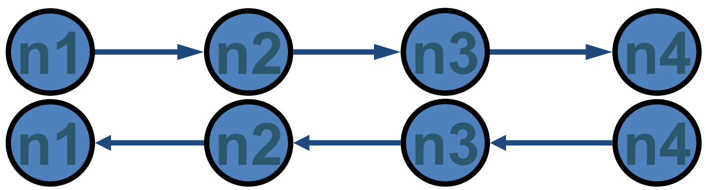


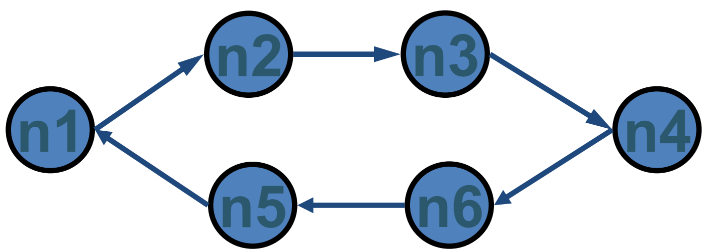
- *Components*
  - Connected subgraphs in a graph
  - Connected graph has 1 component
  - Two disconnected graphs are one social network!!!

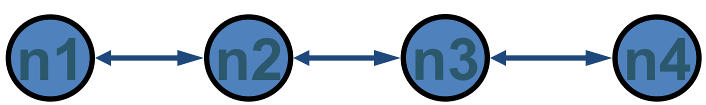


Connected Graphs, Components, Cutpoints and Bridges





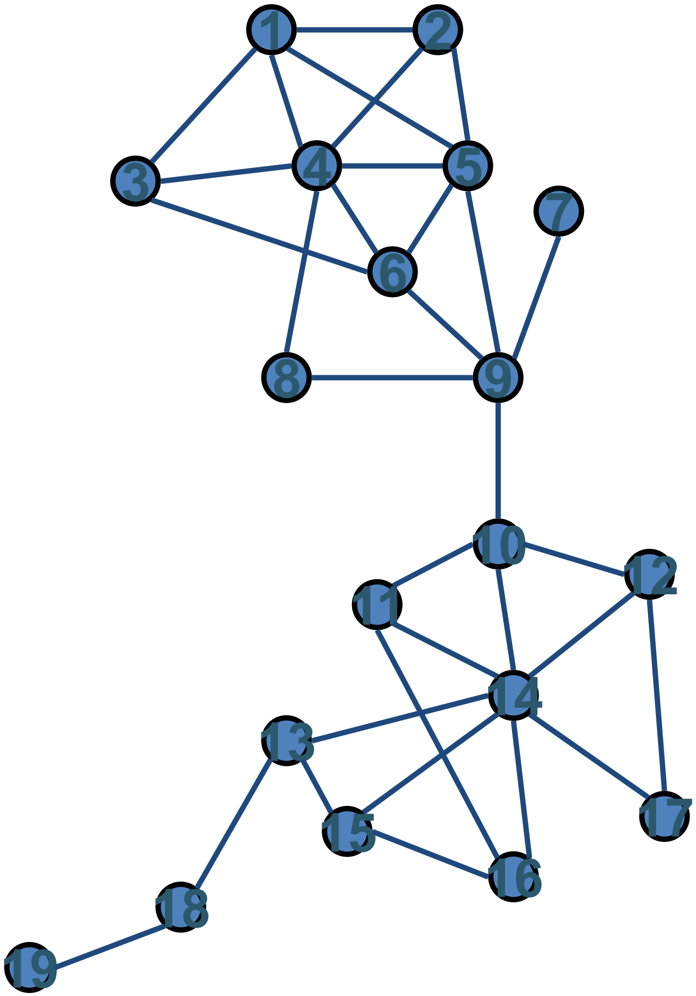




- Connectivity of pairs of nodes and graphs
  - *Weakly connected*
    - Joined by semipath
  - *Unilaterally connected*
    - Path from n_i to n_j or from n_j to n_i
  - *Strongly connected*
    - Path from n_i to n_j *and* from n_j to n_i
    - Path may contain different nodes
  - *Recursively Connected*
    - Nodes are strongly connected and both paths use the same nodes and arcs in reverse order

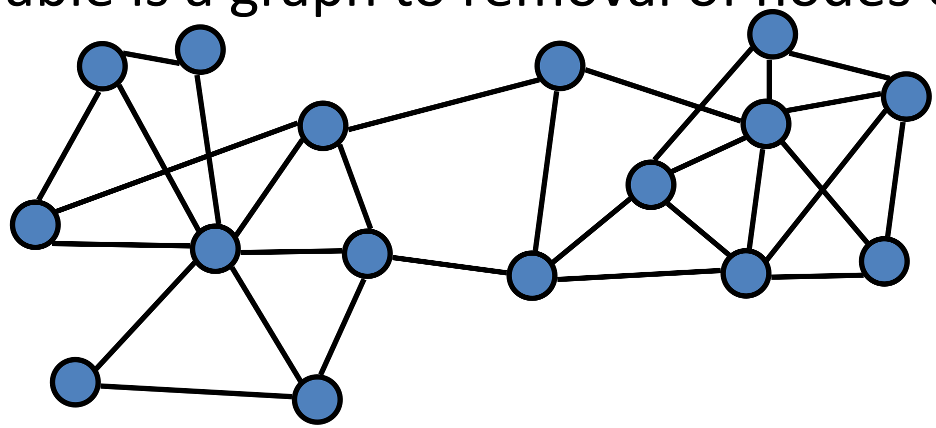
Connected Graphs, Components, Cutpoints and Bridges

- *Cutpoints*
  - number of components in the graph that contain node n_j is fewer than number of components in subgraphs that results from deleting n_j from the graph
- *Cutsets (of size k)*
  - *k*-node cut
- *Bridges / line cuts*
  - Number of components…that contain line l_k



Node- and Line Connectivity

- How vulnerable is a graph to removal of nodes or lines?



Point connectivity / Node connectivity

- Minimum number of *k* for which the graph has a *k*-node cut
- For any value <k the graph is *k-node-connected*

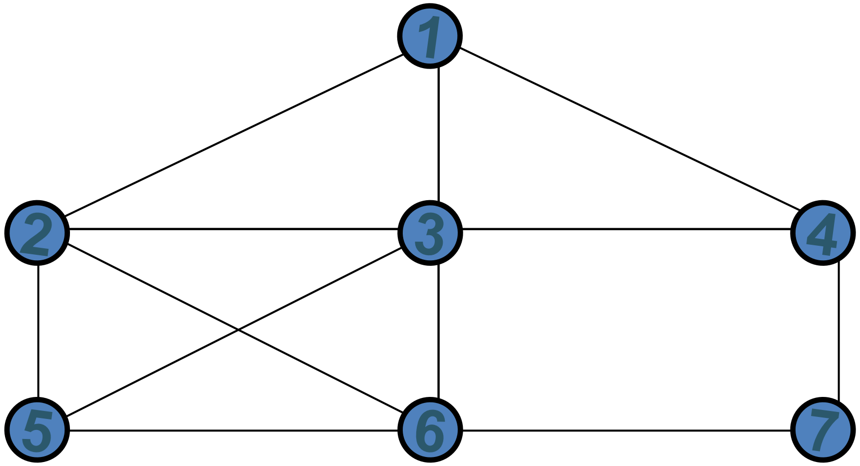
Line connectivity / Edge connectivity

- Minimum number λ for which for which graph has a λ-line cut

Cohesive Subgroups

Cohesive Subgroups, (n-)Cliques, n-Clans, n-Clubs, k-Plexes, k-Cores

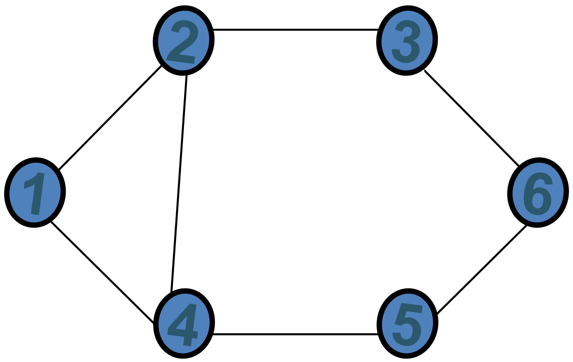
- *Cohesive Subgroup*
  - Subset of actors among there are relatively strong, direct, intense, frequent or positive ties
- *Complete Graph*
  - All nodes are adjacent
- *Clique*
  - Maximal complete subgraph of three or more nodes
  - Cliques can overlap
  - {1, 2, 3}
  - {1, 3, 4}
  - {2, 3, 5, 6}



Cohesive Subgroups, (n-)Cliques, n-Clans, n-Clubs, k-Plexes, k-Cores

- *n-clique*
  - maximal subgraph in which d(i,j) ≤ n for all n_i, n_j
  - 2: cliques: {2, 3, 4, 5, 6} and {1, 2, 3, 4, 5}
  - intermediaries in geodesics do not have to be n-clique members themselves!

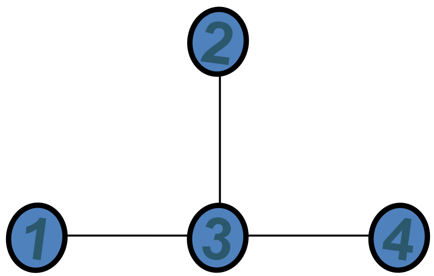
- *n-clan*
  - *n-clique* in which the d(i,j) ≤ n for the subgraph of all nodes in the n-clique
  - 2-clan: {2, 3, 4, 5, 6}

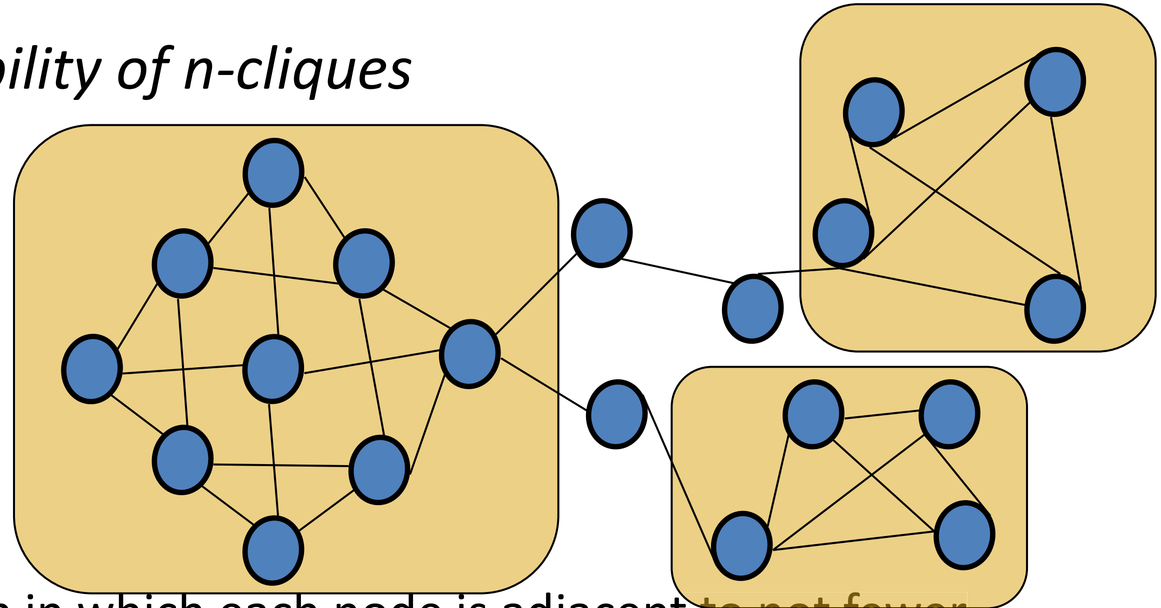


- *n-club*
  - maximal subgraph of diameter n
  - 2-clubs: {1, 2, 3, 4} ; {1, 2, 3, 5} and {2, 3, 4, 5, 6}

Cohesive Subgroups, (n-)Cliques, n-Clans, n-Clubs, k-Plexes, k-Cores

- *Problem: vulnerability of n-cliques*





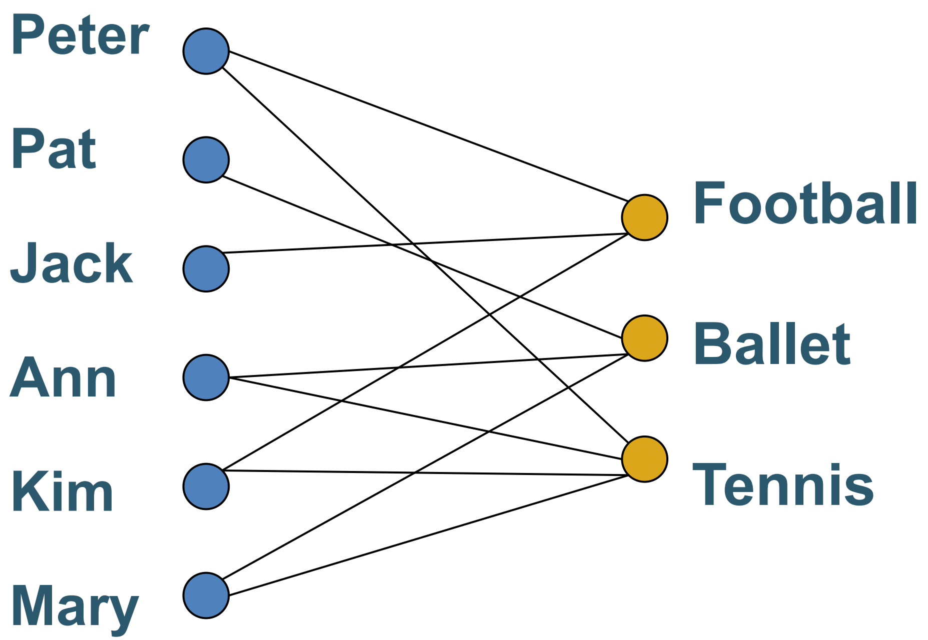
- *k-plexes*
  - maximal subgraph in which each node is adjacent to not fewer than g_s-k nodes („maximal“: no other nodes in subgraph that also have $d_s(i) \geq (g_s - k)$]
- *k-cores*
  - subgraph in which each node is adjacent to at least k other nodes in the subgraph

Analyzing Affiliation Networks

Affiliation Matrix, Bipartite Graph and Hypergraph, Rate of Participation, Size of Events

Two-mode network / affiliation network / membership network / hypernetwork

- nodes can be partitioned in two subsets
  - N (for example g persons)
  - M (for example h clubs)
- depicted in *Bipartite Graph*
- lines between nodes belonging to different subsets



Affiliation Matrix, Bipartite Graph and Hypergraph, Rate of Participation, Size of Events

Affiliation Matrix (Incidence Matrix)

- Connections among members of one of the modes based on linkages established through second mode
- g actors, h events
- $A = \{a_{ij}\}$ $(g \times h)$

Event

Actor		Football	Ballet	Tennis	rate of participation
	Peter	1		1	1
	Pat		1		1
	Jack	1			1
	Ann		1	1	2
	Kim	1		1	2
	Mary		1	1	2
size of event		3	3	4	

Affiliation Matrix, Bipartite Graph and Hypergraph, Rate of Participation, Size of Events

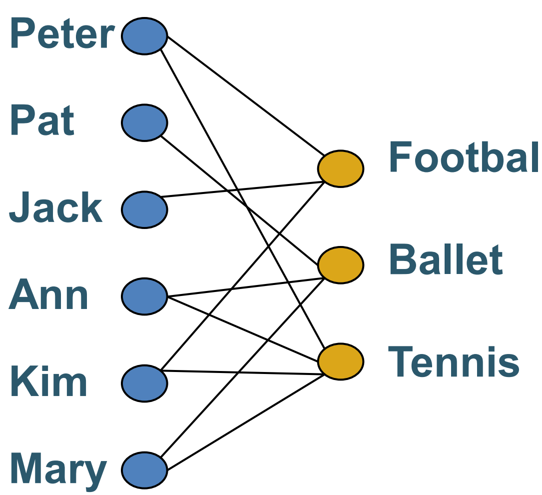
- Sociomatrix [(g+h) × (g+h)]

	Peter	Pat	Jack	Ann	Kim	Mary	Football	Ballet	Tennis
Peter	-	0	0	0	0	0	1	0	1
Pat	0	-	0	0	0	0	0	1	0
Jack	0	0	-	0	0	0	1	0	0
Ann	0	0	0	-	0	0	0	1	1
Kim	0	0	0	0	-	0	1	0	1
Mary	0	0	0	0	0	-	0	1	1
Football	1	0	1	0	1	0	-	0	0
Ballet	0	1	0	1	0	1	0	-	0
Tennis	1	0	0	1	1	1	0	0	-

Affiliation Matrix, Bipartite Graph and Hypergraph, Rate of Participation, Size of Events

- *Homogenous* pairs and *heterogenous* pairs
- X_r^N (g×g), X_r^M (h×h), $X_r^{N,M}$(g×h), $X_r^{N,M}$(h×g)
  - One-mode sociomatrices X^N [and X^M]
    - rows, colums: actors [events];
    - x_{ij}: *co-membership* [number of actors in both events] (main diagonal meaningful, e.g. total events attended by an actor)

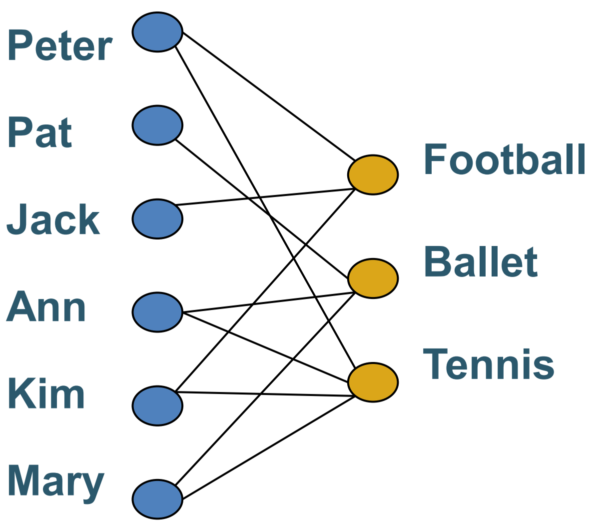
	Peter	Pat	Jack	Ann	Kim	Mary
Peter	2	0	1	1	2	1
Pat	0	1	0	1	0	1
Jack	1	0	1	0	1	0
Ann	1	0	0	2	1	1
Kim	2	0	1	1	2	1
Mary	1	0	0	1	1	2



Affiliation Matrix, Bipartite Graph and Hypergraph, Rate of Participation, Size of Events

- *Event Overlap / Interlocking Matrix*

	Football	Ballet	Tennis
Football	3	0	2
Ballet	0	3	1
Tennis	2	1	4



Cohesive Subsets of Actors or Events

- *clique at level c (cf. also k-plexes, n-cliques etc.)*
  - subgraph in which all pairs of events share at least c members
- *connected at level q*
  - subset in which all actors in the path are co-members of at least q+1 events

...HOLDOUT I...

When is Which Centrality Measure Appropriate?

Source: *Borgatti, Stephen P. (2005) Centrality and Network Flow, Social Networks 27, p. 55-71*

Assumptions of Centrality Measures

- Which things flow through a network and how do they flow?

	Transfer	Serial	Parallel
Walks	Money exchange	Emotional support	Attitude influencing
Trails	Used Book	Gossip	E-mail broadcast
Paths	Mooch	Viral infection	Internet name-server
Geodesics	Package Delivery	Mitotic reproduction	<no process>

Source: *Borgatti, Stephen P. (2005) Centrality and Network Flow, Social Networks 27, p. 55-71*

Assumptions of Centrality Measures

- Example: Betweenness Centrality
  - Information travels along the shortest route

	Transfer	Serial	Parallel
Walks	Random Walk Betweenness	?	Closeness Degree Eigenvector
Trails	?	?	Closeness Degree
Paths	?	?	Closeness Degree
Geodesics	Closeness Betweenness	Closeness	?

Adequacy of Centrality Measures

	Transfer	Serial	Parallel
Walks	Money exchange	Emotional support	Attitude influencing
Trails	Used Book	Gossip	E-mail broadcast
Paths	Mooch	Viral infection	Internet name-server
Geodesics	Package Delivery	Mitotic reproduction	<no process>

Source: *Borgatti, Stephen P. (2005) Centrality and Network Flow, Social Networks 27, p. 55-71*

How to Calculate Geodesic Distance Matrices?

From Adjacency Matrices to (Geodesic) Distance Matrices I – (Reachability)

Repetition: Matrix Multiplication

- XY = Z

X

3	0	2
1	4	2

g × h

Y

2	3
1	4
4	2

h × k

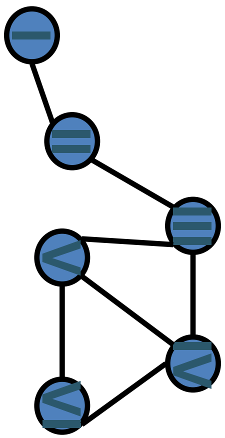
Z

14	13
14	23

g × k

- $z_{ij} = \sum_{n=1}^{h} x_{in} \, y_{nj}$

From Adjacency Matrices to (Geodesic) Distance Matrices II – (Reachability)



	I	II	III	IV	V	VI
I	0	1	0	0	0	0
II	1	0	1	0	0	0
III	0	1	0	1	1	0
IV	0	0	1	0	1	1
V	0	0	1	1	0	1
VI	0	0	0	1	1	0

X

	I	II	III	IV	V	VI
I	0	1	0	0	0	0
II	1	0	1	0	0	0
III	0	1	0	1	1	0
IV	0	0	1	0	1	1
V	0	0	1	1	0	1
VI	0	0	0	1	1	0

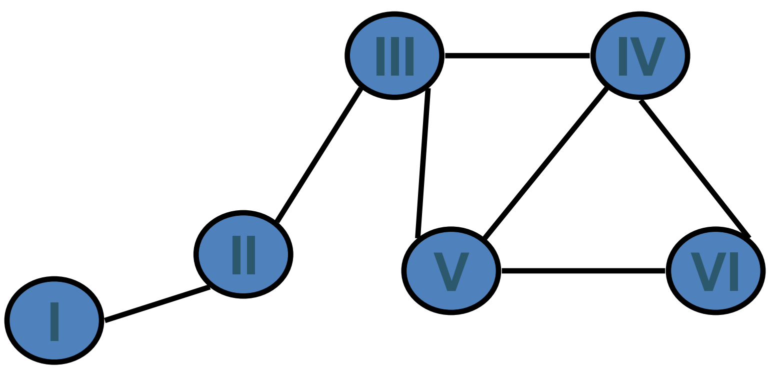
Power Matrix: Multiplying adjacency matrices

	I	II	III	IV	V	VI
I	1	0	1	0	0	0
II	0	2	0	1	1	0
III	1	0	3	1	1	2
IV	0	1	1	3	2	1
V	0	1	1	2	3	1
VI	0	0	2	1	1	2

- **$x_{ik}x_{kj} = 1$ only if lines (n_i, n_k) and (n_k, n_j) are present, i.e. $X^{[2,3,4]}$ counts the number of walks $(n_i n_k n_j)$ of length 1 [2,3,4] between nodes n_i and n_j**

From Adjacency Matrices to (Geodesic) Distance Matrices II – (Reachability)

- $x_{ij}>0$?
  ➔ two nodes can be connected by paths of length ≤ (g-1)
- Calculate $X^{[\Sigma]} = X + X^2 + X^3 + \ldots + X^{g-1}$
- $X^{[\Sigma]}$ shows total number of walks from n_i to n_j

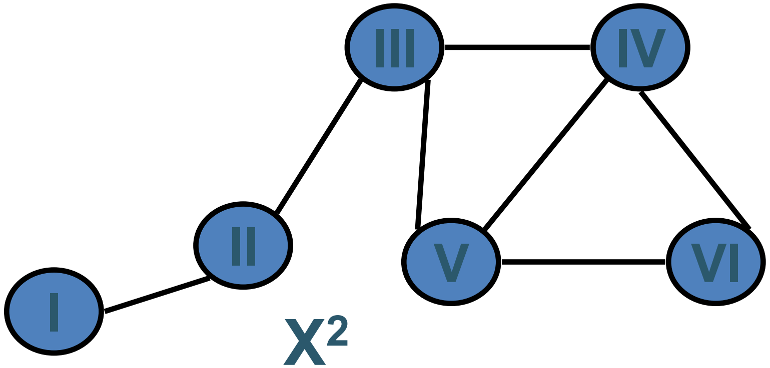


X^2

	I	II	III	IV	V	VI
I	1	0	1	0	0	0
II	0	2	0	1	1	0
III	1	0	3	1	1	2
IV	0	1	1	3	2	1
V	0	1	1	2	3	1
VI	0	0	2	1	1	2

From Graphs to (Geodesic Distance)-Matrices (Reachability) – Geodesic Distance

- observer power matrices
- first power p for which the (i,j) element is non-zero gives the shortest path
- $d(i,j) = \min_p x_{ij}^{[p]} > 0$



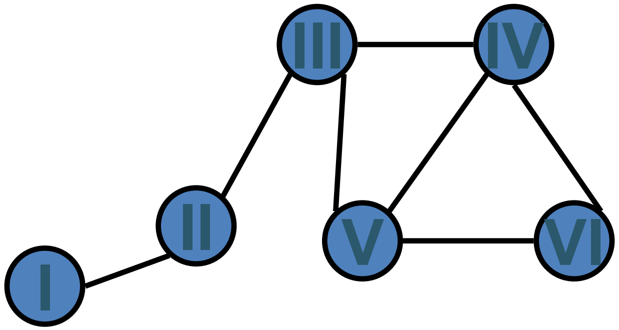
X

	I	II	III	IV	V	VI
I	0	1	0	0	0	0
II	1	0	1	0	0	0
III	0	1	0	1	1	0
IV	0	0	1	0	1	1
V	0	0	1	1	0	1
VI	0	0	0	1	1	0

X^2

	I	II	III	IV	V	VI
I	1	0	1	0	0	0
II	0	2	0	1	1	0
III	1	0	3	1	1	2
IV	0	1	1	3	2	1
V	0	1	1	2	3	1
VI	0	0	2	1	1	2

From Graphs to (Geodesic Distance)-Matrices (Reachability) – Geodesic Distance



Binary, undirected

	I	II	III	IV	V	VI
I	-	1	2	3	3	4
II	1	-	1	2	2	3
III	2	1	-	1	1	2
IV	3	2	1	-	1	1
V	3	2	1	1	-	1
VI	4	3	2	1	1	-