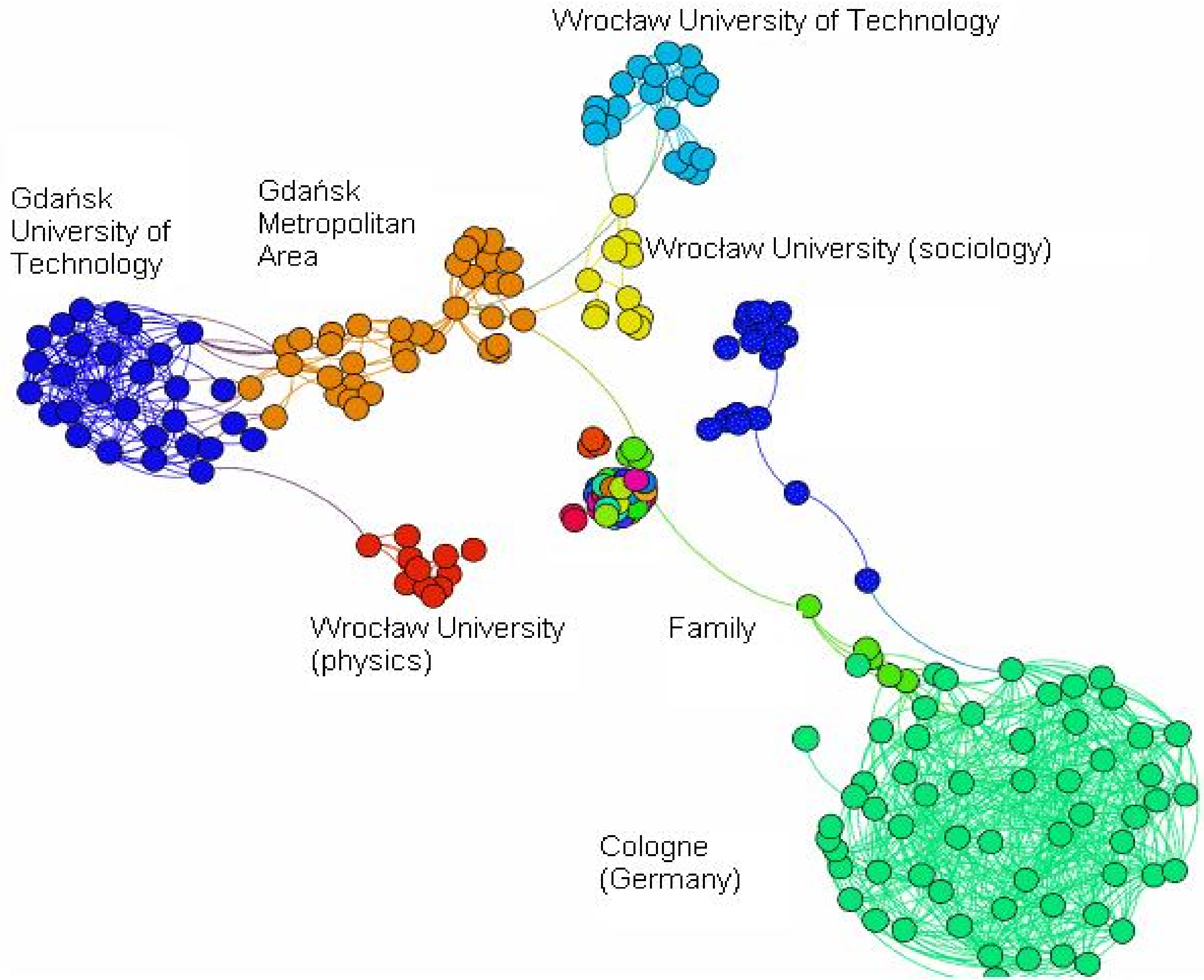
Crash course on network analysis

Joanna Byszuk & Jeremi Ochab

DHSI 2024, “DIY Computational Text Analysis with R”

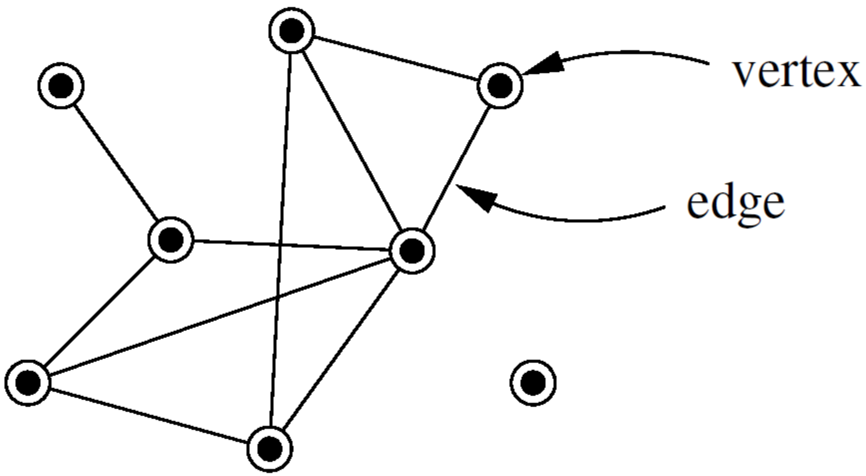
Outline

1. Types of graphs & representation
2. Examples
3. Network structures, centrality, community detection


Wrocław University of Technology
Gdańsk University of Technology
Gdańsk Metropolitan Area
Wrocław University (sociology)
Wrocław University (physics)
Family
Cologne (Germany)

NETWORK TYPES
AND REPRESENTATIONS

Types of graphs/networks

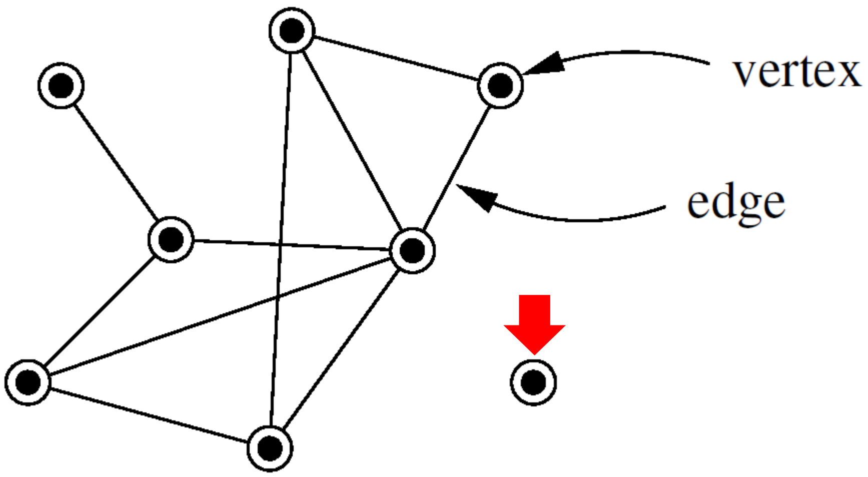


Entities (vertices/nodes)

Relations (edges/links)

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Types of graphs



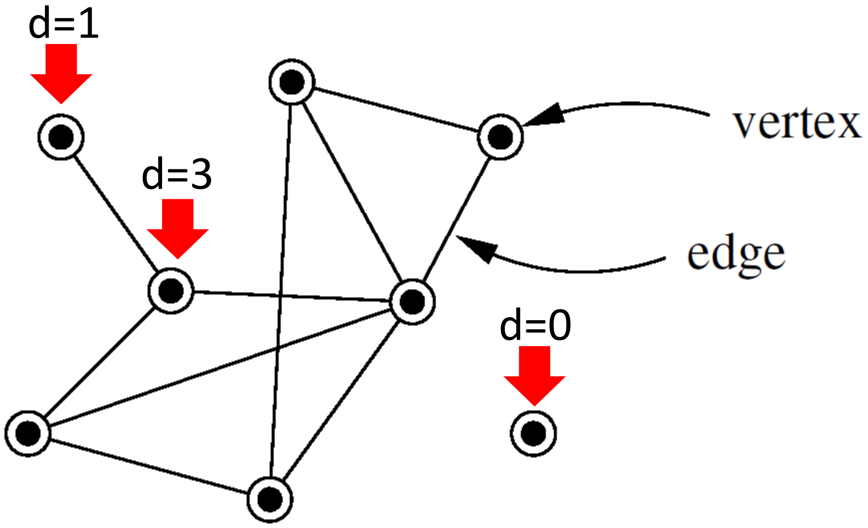
Entities (vertices/nodes)

Relations (edges/links)

Component

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Types of graphs



Entities (vertices/nodes)

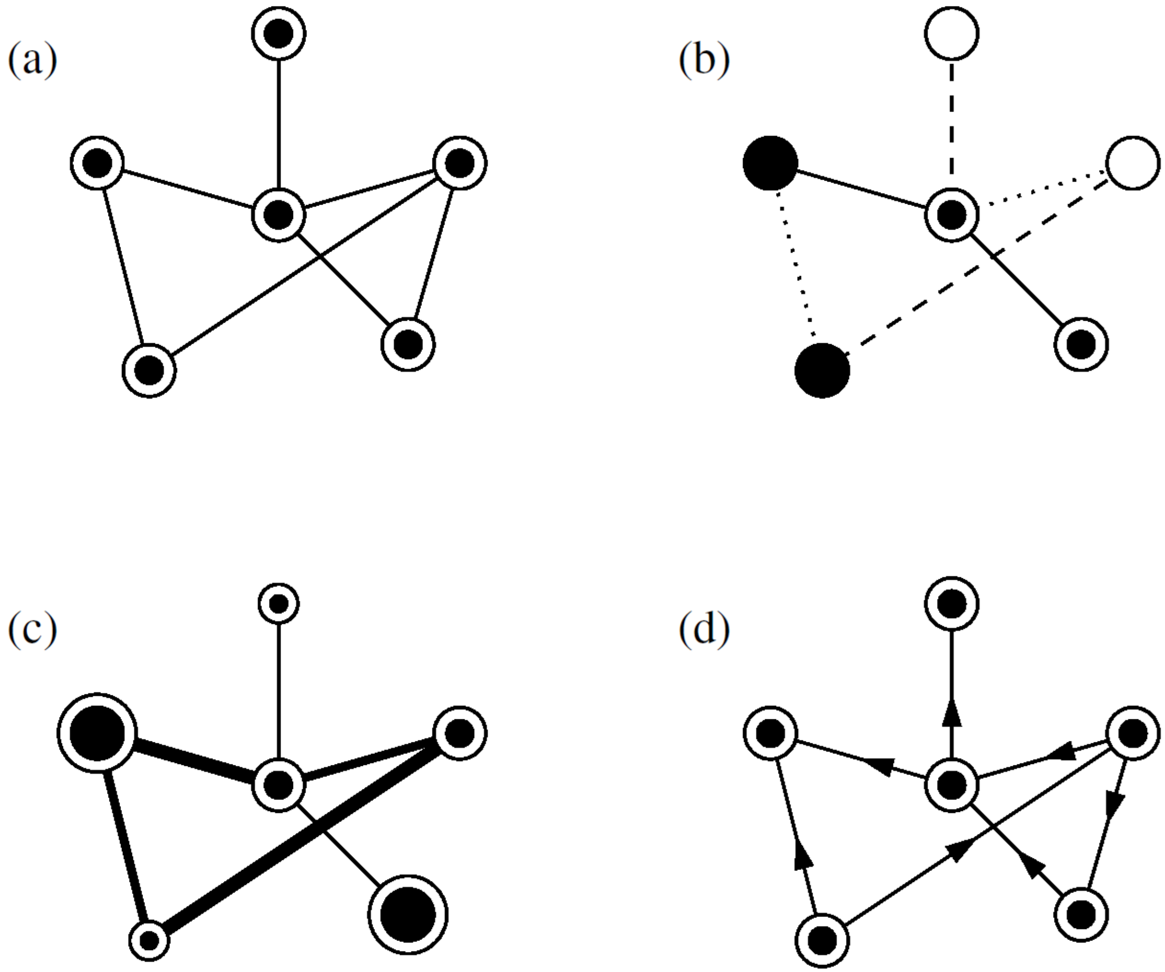
Relations (edges/links)

Component

Degree

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Types of graphs

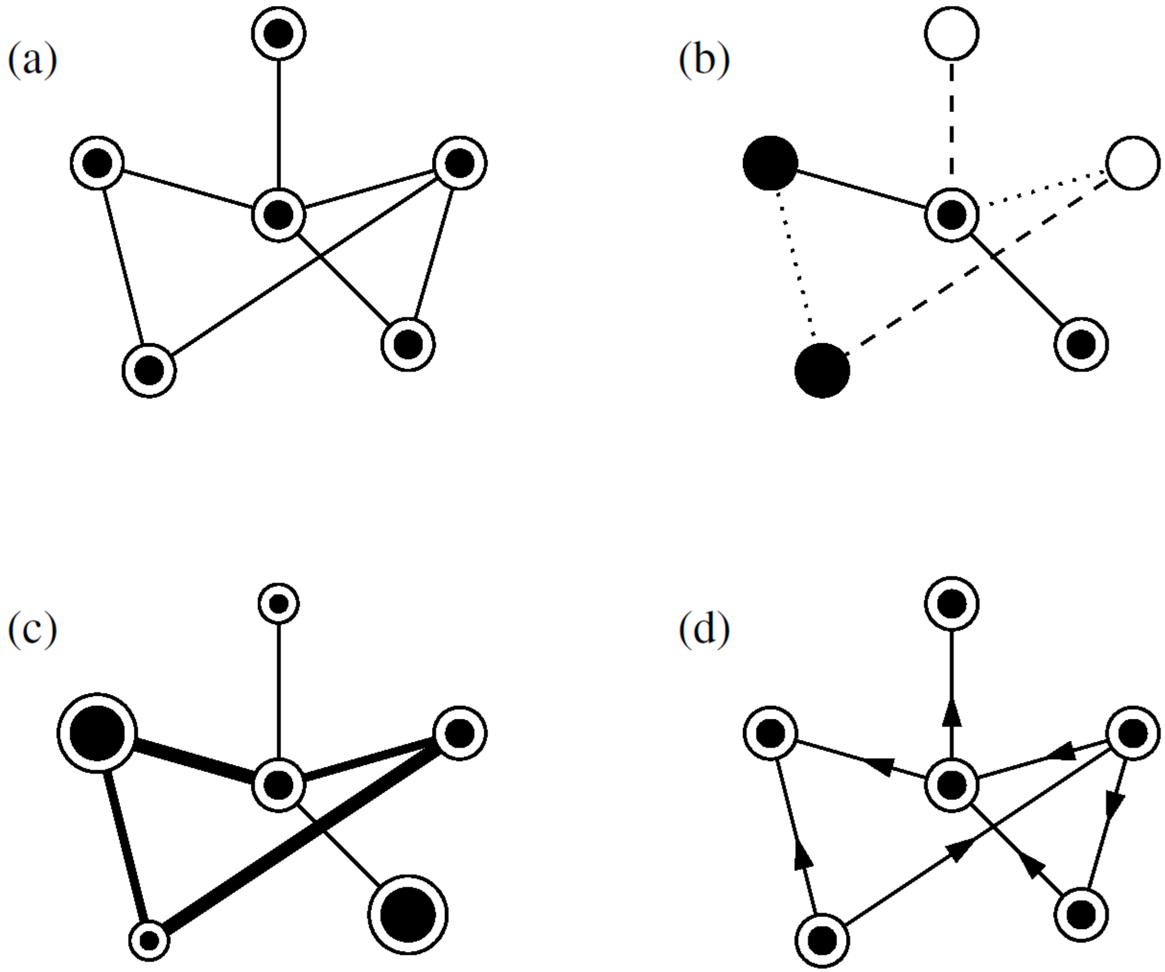


a) simple (undirected)
b) multi-edge, multi-vertex type

a) weighted
b) directed

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Types of graphs



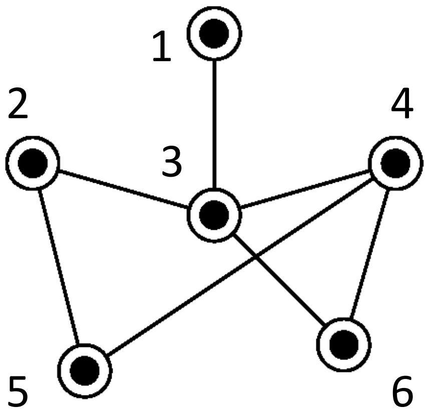
Networks:

- social
- information
- technological
- biological

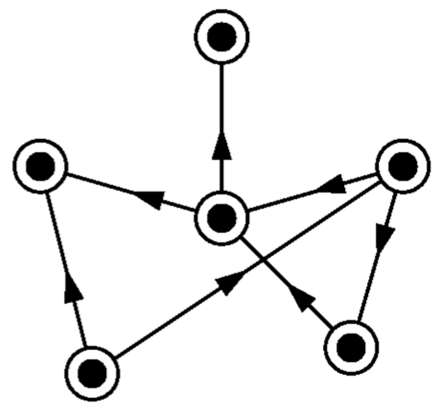
MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Representations: adjacency matrix

(a)



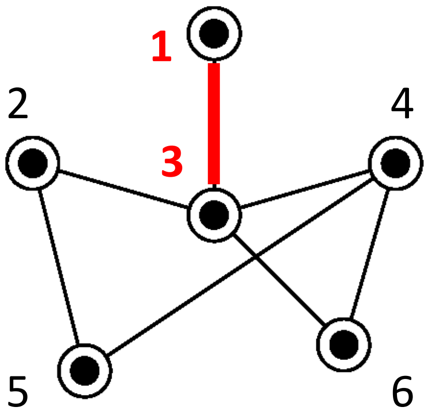
(d)



MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

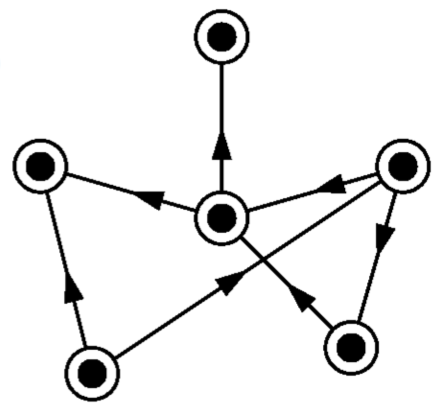
Representations: adjacency matrix

(a)



		1			
1					

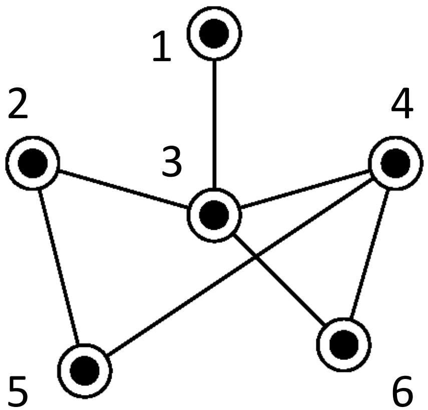
(d)



MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

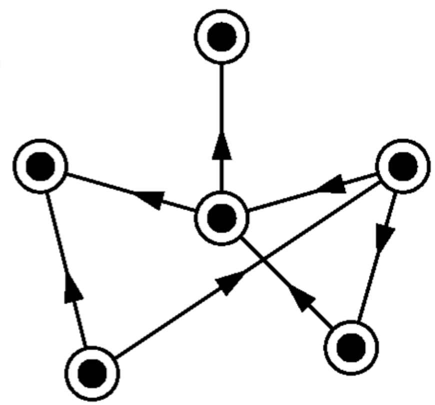
Representations: adjacency matrix

(a)



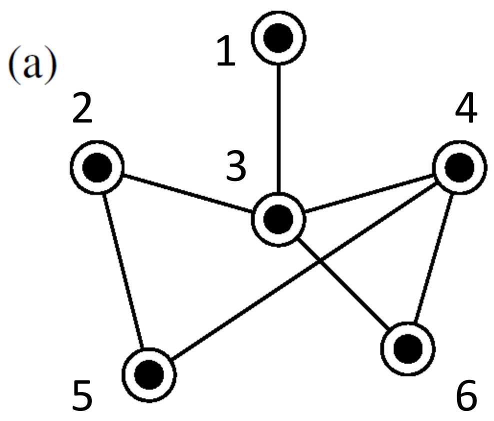
		1			
		1		1	
1	1		1		1
		1		1	1
	1		1		
		1	1		

(d)

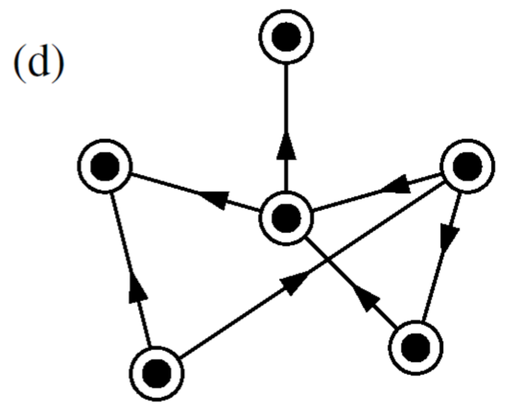


MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Representations: adjacency matrix

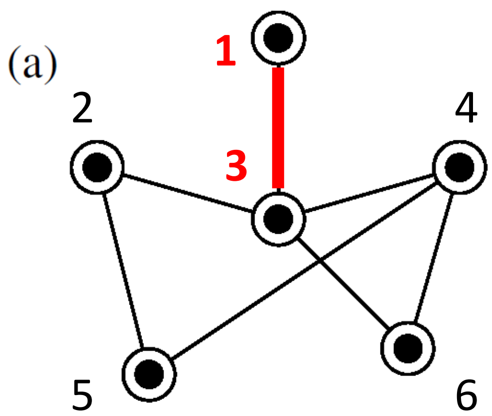


0	0	1	0	0	0
0	0	1	0	1	0
1	1	0	1	0	1
0	0	1	0	1	1
0	1	0	1	0	0
0	0	1	1	0	0

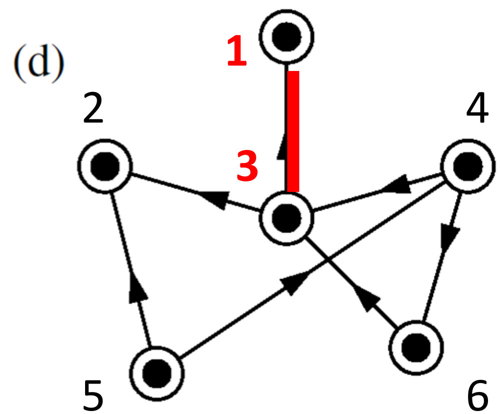


MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Representations: adjacency matrix



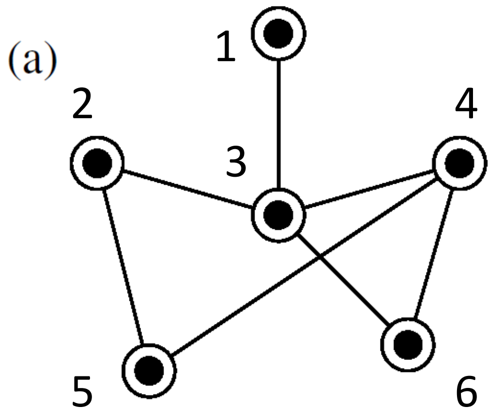
		1			
		1		1	
1	1		1		1
		1		1	1
	1		1		
		1	1		



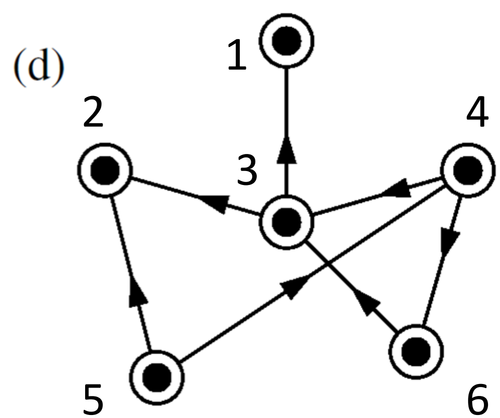
		0			
		1		1	
1	1		1		1
		1		1	1
	1		1		
		1	1		

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Representations: adjacency matrix



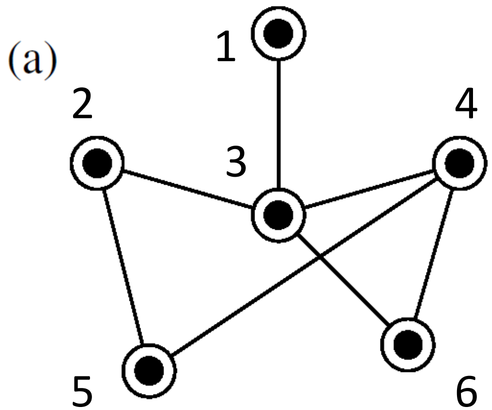
		1			
		1		1	
1	1		1		1
		1		1	1
	1		1		
		1	1		



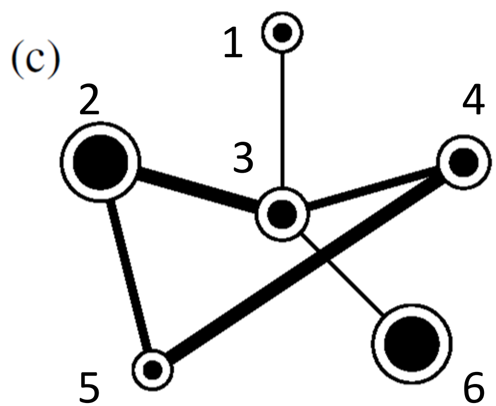
1	1				
		1			1
	1		1		
		1			

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Representations: adjacency matrix



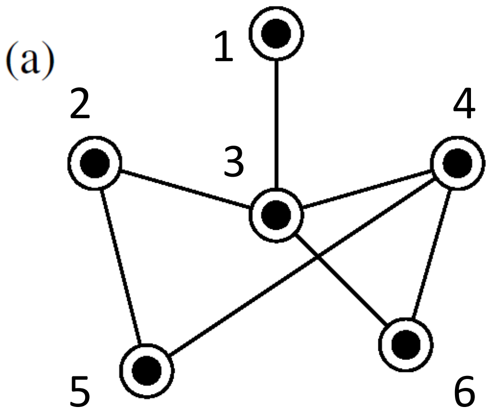
		1			
		1		1	
1	1		1		1
		1		1	1
	1		1		
		1	1		



		1			
		3		2	
1	3		2		1
		2		3	1
	2		3		
		1	1		

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Representations: edge list



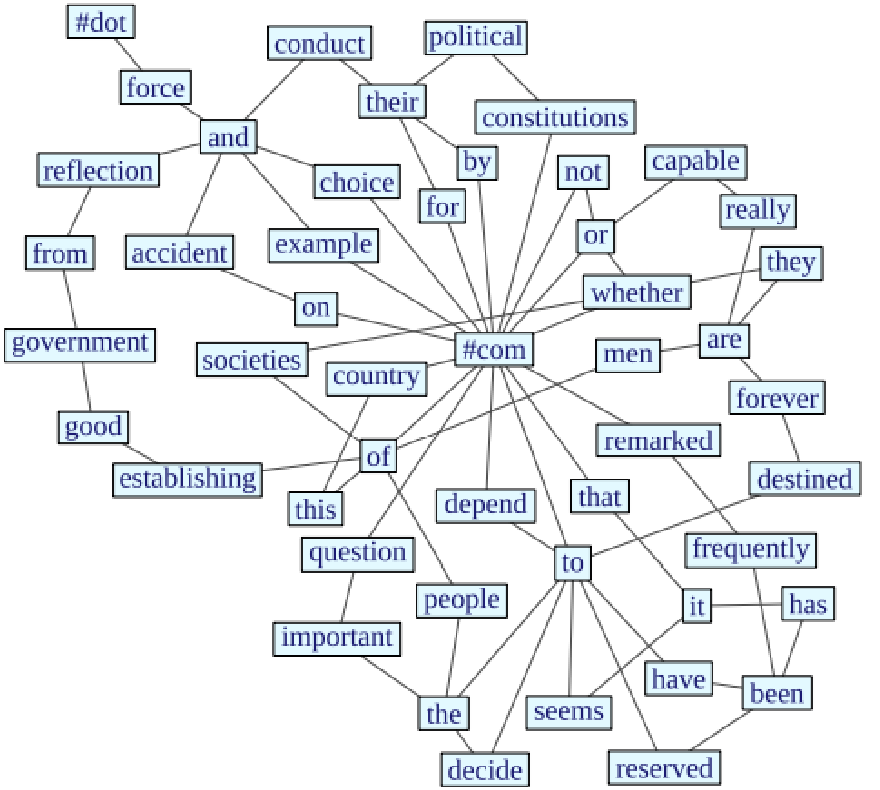
	Nodes					
Nodes			1			
			1		1	
	1	1		1		1
			1		1	1
		1		1		
			1	1		

Edges
1-3
2-3
2-5
3-4
3-6
4-5
4-6

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

NETWORK EXAMPLES

Co-occurence networks



Stanisz T, Kwapień J, Drożdż S (2019). Linguistic data mining with complex networks: A stylometric-oriented approach, *Information Sciences,* 482, 301-320.

Newman ME (2006). Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev .E 74*, 036104.

Dependency networks

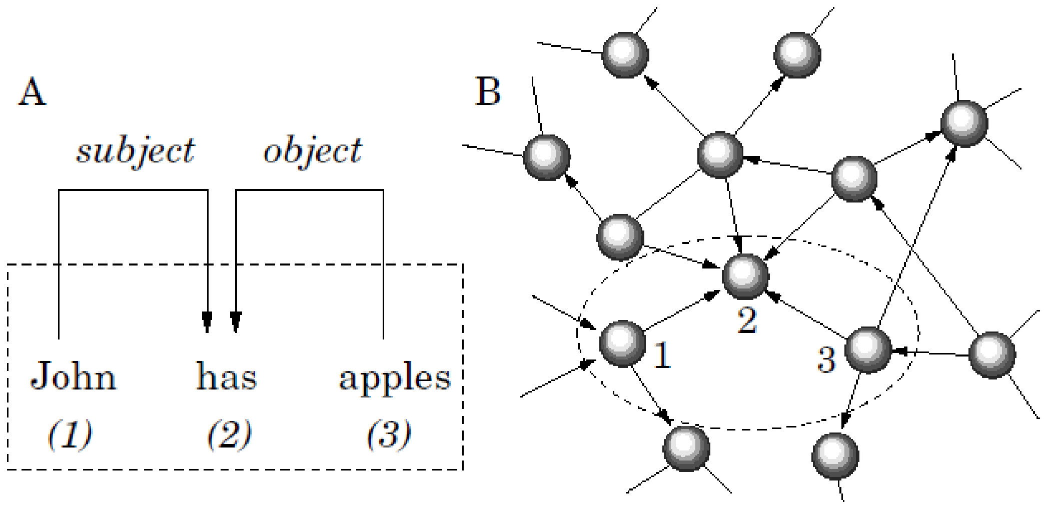
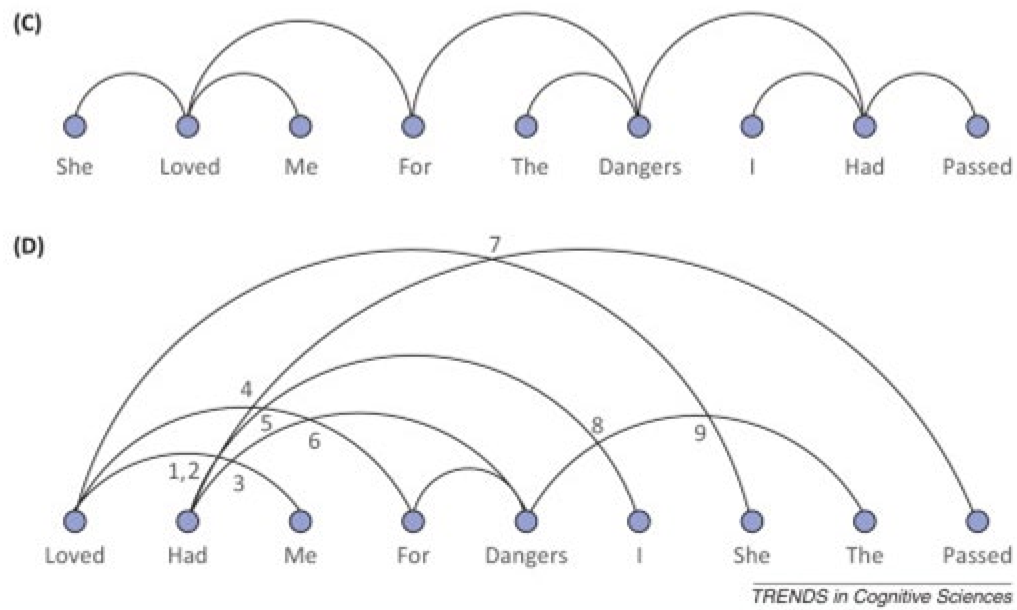


FIG. 1: A) The syntactic structure of a simple sentence. Words are the vertices and the syntactic dependencies are the edges of the graph. The proper noun 'John' and the verb 'has' are syntactically dependent in this sample sentence. 'John' is *modifier* of the verb 'has', which is its *head*. Similarly, the action of 'has' is modified by its object 'apples'. Here we assume the graph oriented with edges pointing from a *modifier* to its *head*. B) Mapping the syntactic dependency structure of the sentence into a global syntactic dependency network.

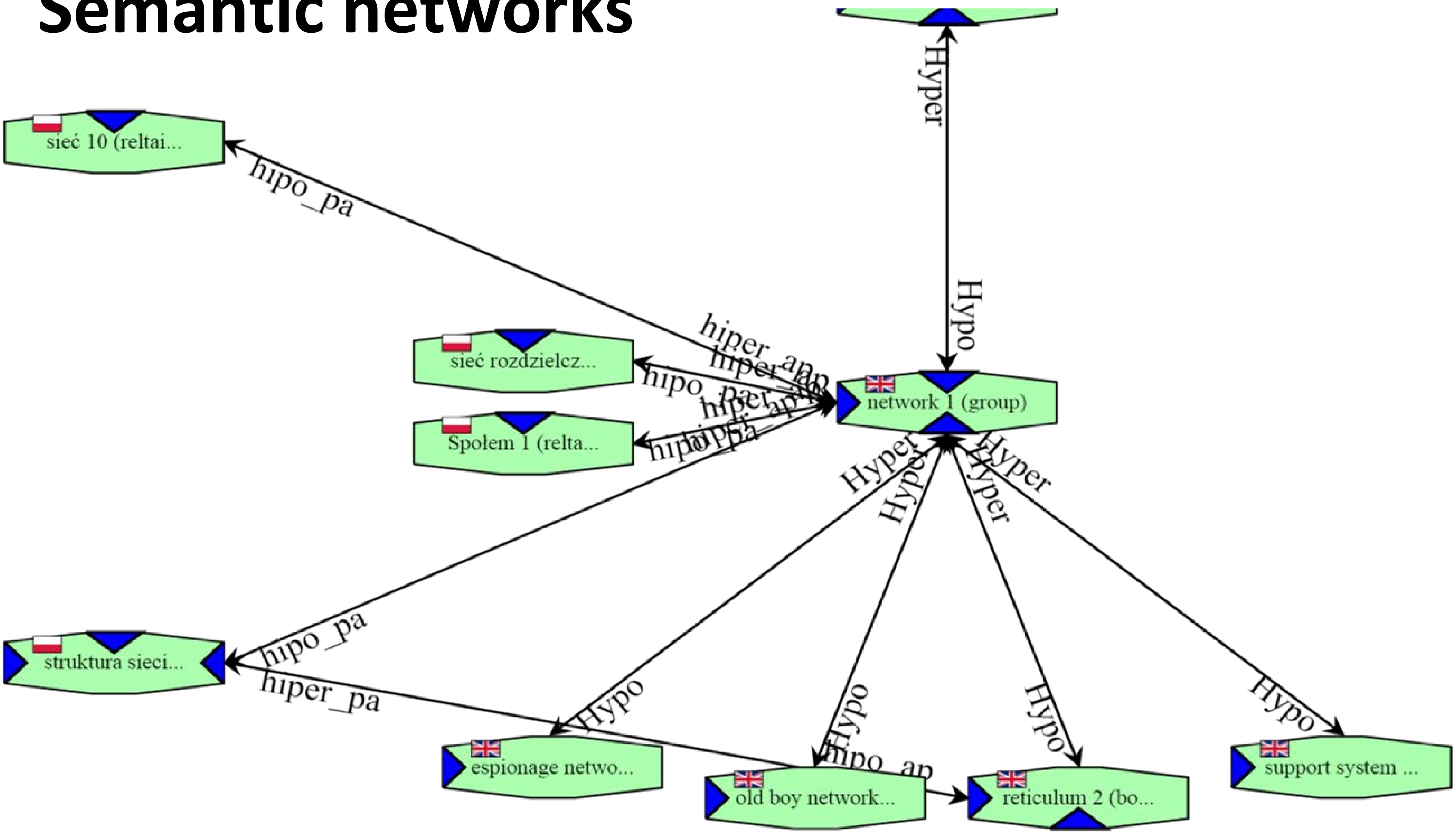
Ferrer-i-Cancho R, et al. (2007). Spectral methods cluster words of the same class in a syntactic dependency network. *International Journal of Bifurcation and Chaos*, *17*(07), 2453-2463.

Dependency networks



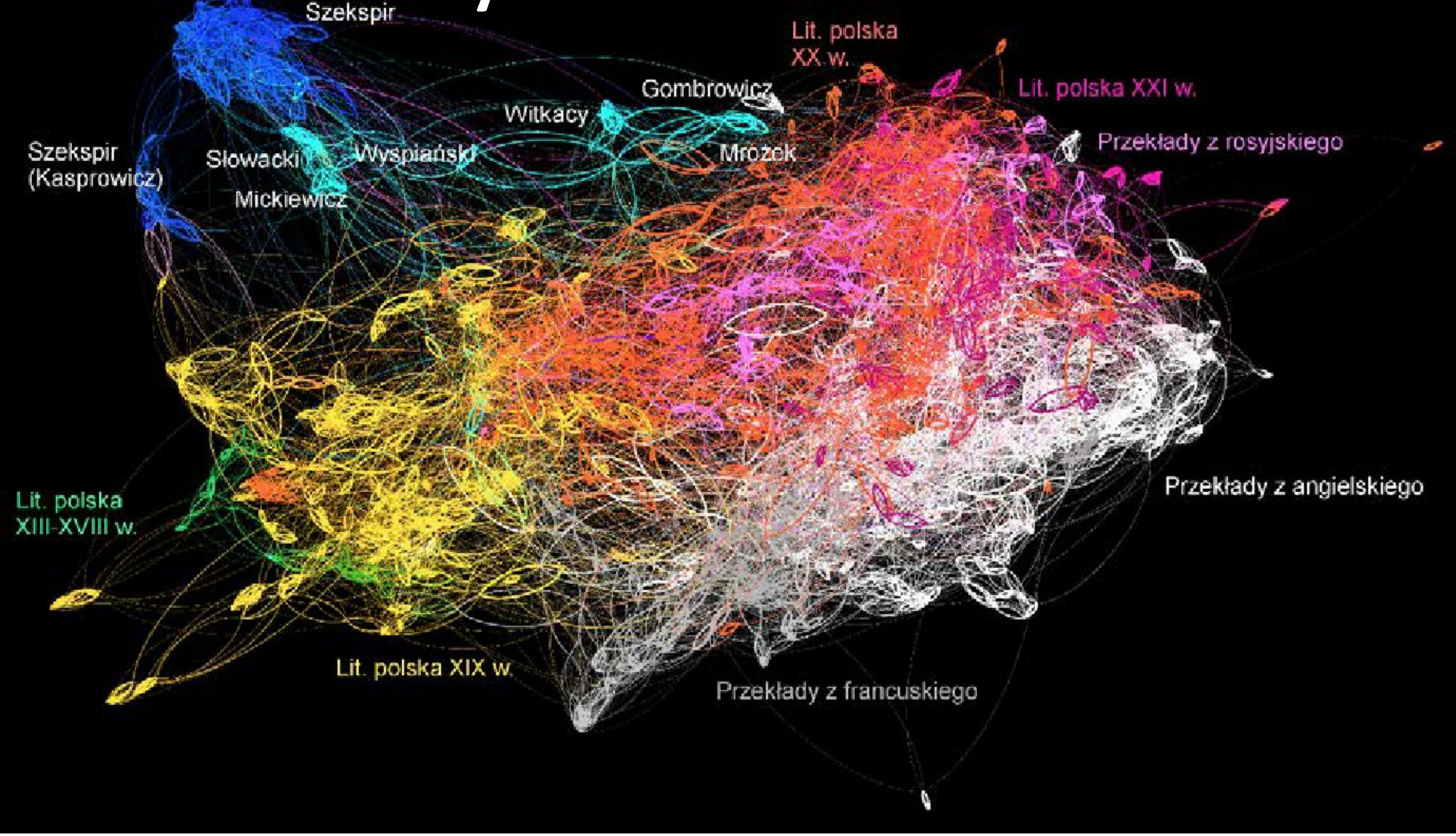
Baronchelli A, Ferrer-i-Cancho R, et al. (2013) Networks in Cognitive Science. Trends in Cognitive Sciences 17, pp. 348-360.

Semantic networks



Piasecki M, Maziarz M, Marcińczuk M, et al. (2010) WordNet, CLARIN-PL digital repository, http://hdl.handle.net/11321/29.
http://plwordnet.pwr.wroc.pl

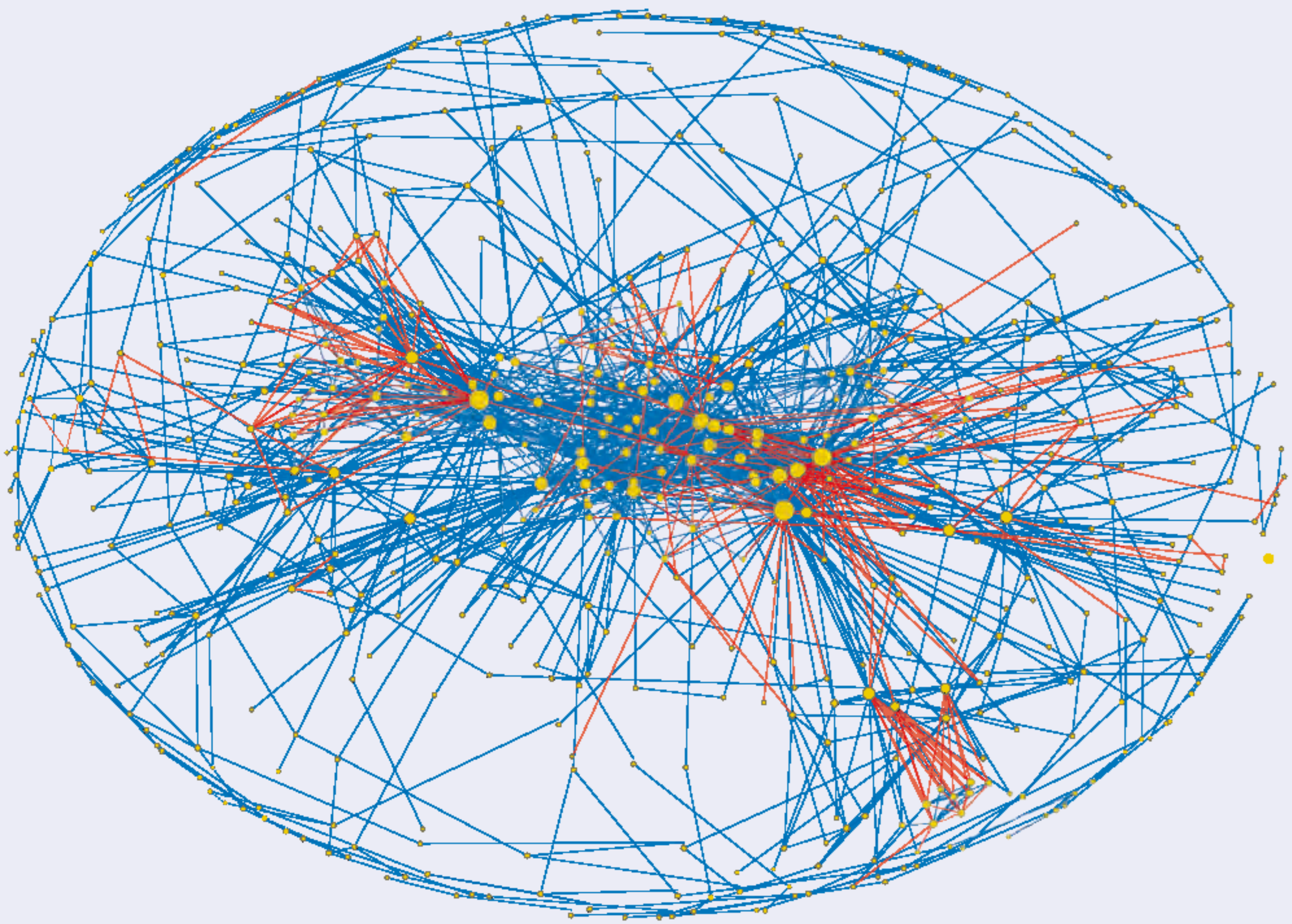
Text similarity networks



Rybicki J (2017) Drugi rzut oka na stylometryczną mapę literatury polskiej. Forum Poetyki 10: 6-21
Eder, M. (2015). *Visualization in stylometry: cluster analysis using networks*. Digital Scholarship in the Humanities, 30

Literary character networks

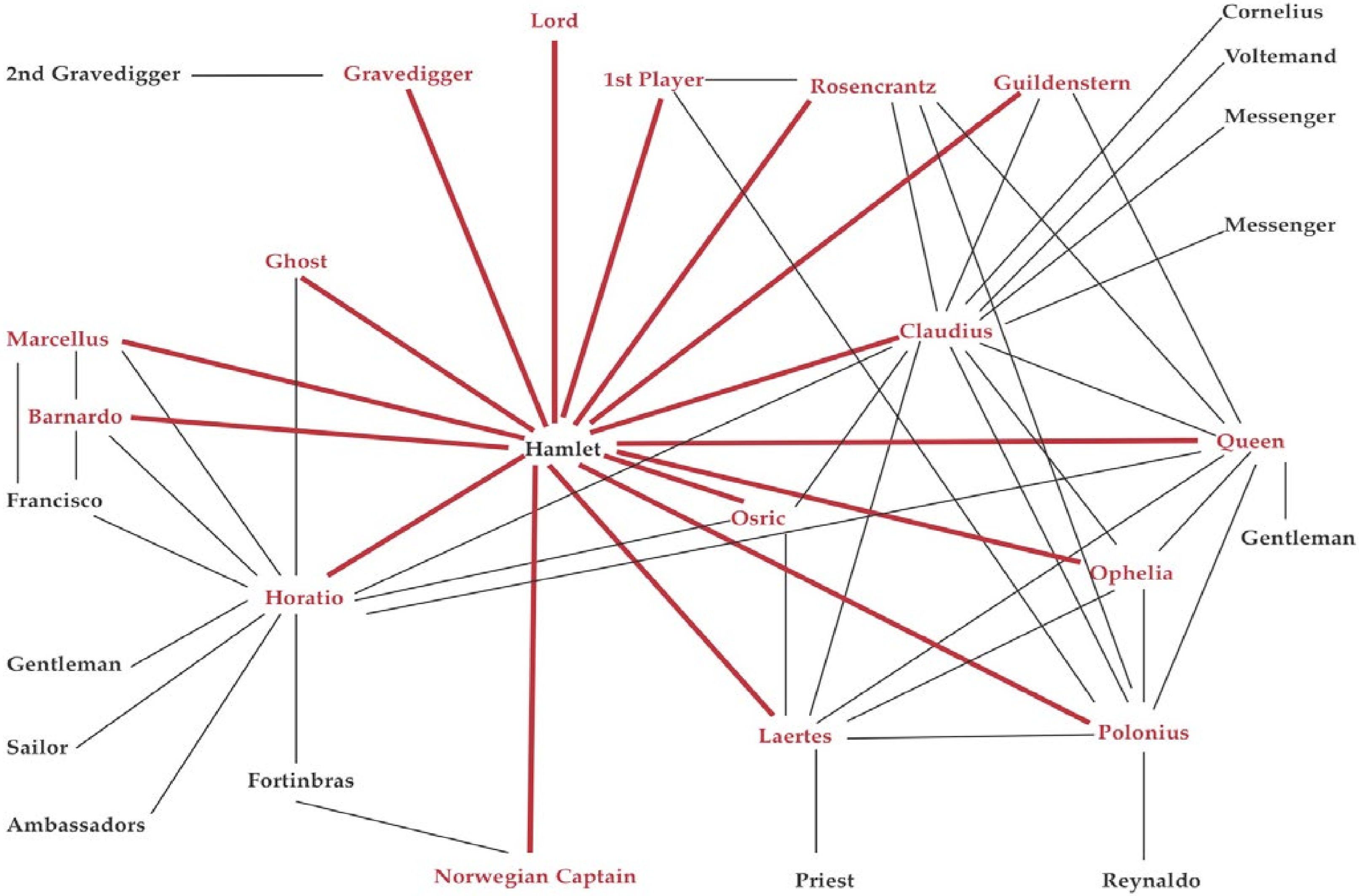
1 The complex network of Njál



Nodes (yellow) represent characters appearing in the narrative of *Njáls saga* and the lines between these nodes, known as "edges", represent relationships between them. Positive (friendly) relationships are shown in blue and negative (hostile) ones are shown in red.

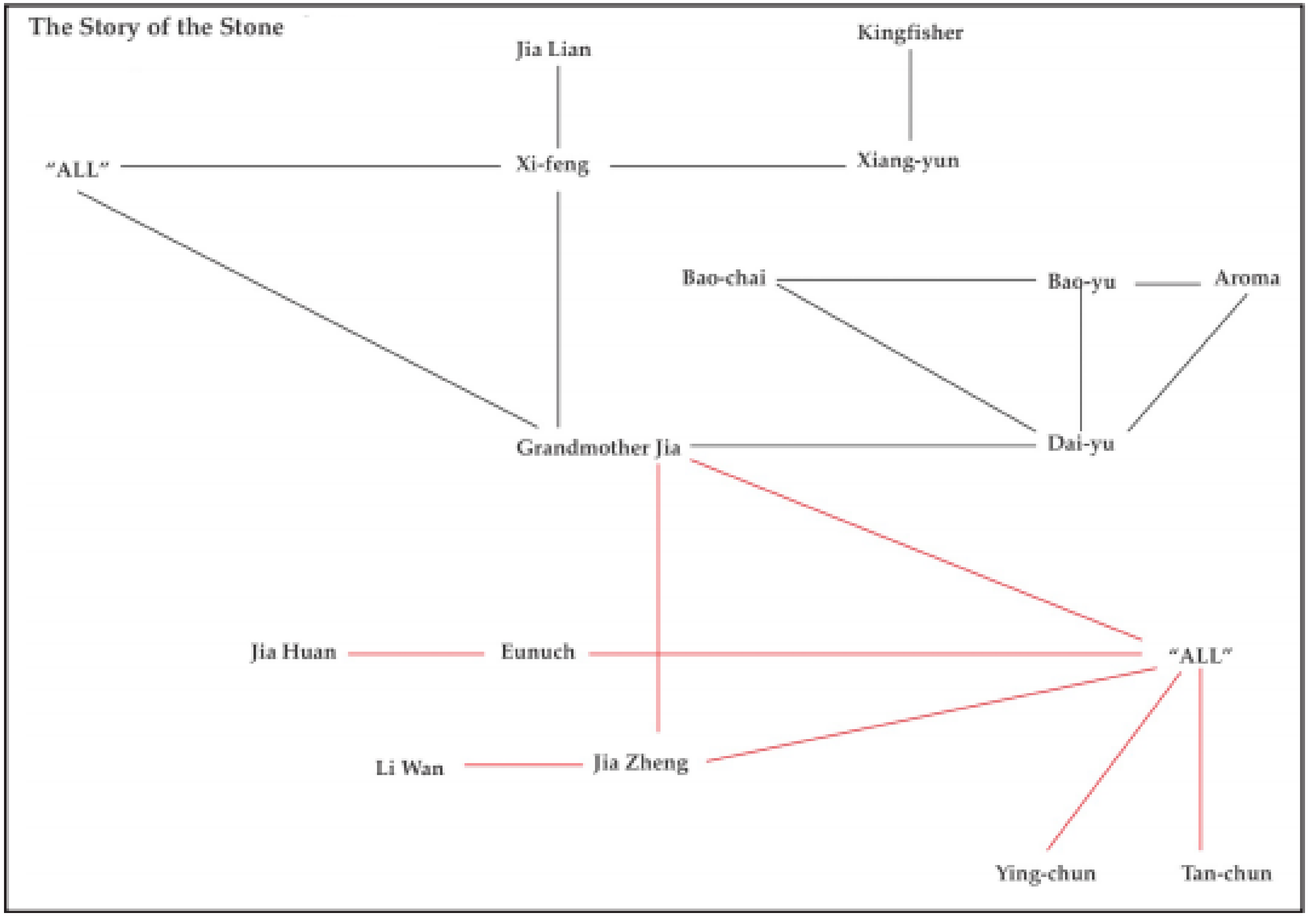
Kenna, R. and MacCarron, P. (2016) Maths meets myths. Physics World, volume 29 (6): 22-27

Literary character networks



Moretti F. 'Network Theory, Plot Analysis'. Web: (http://litlab.stanford.edu/LiteraryLabPamphlet2A.Text.pdf)

Literary character networks



Moretti F. 'Network Theory, Plot Analysis'. Web: (http://litlab.stanford.edu/LiteraryLabPamphlet2A.Text.pdf)

Conversational networks

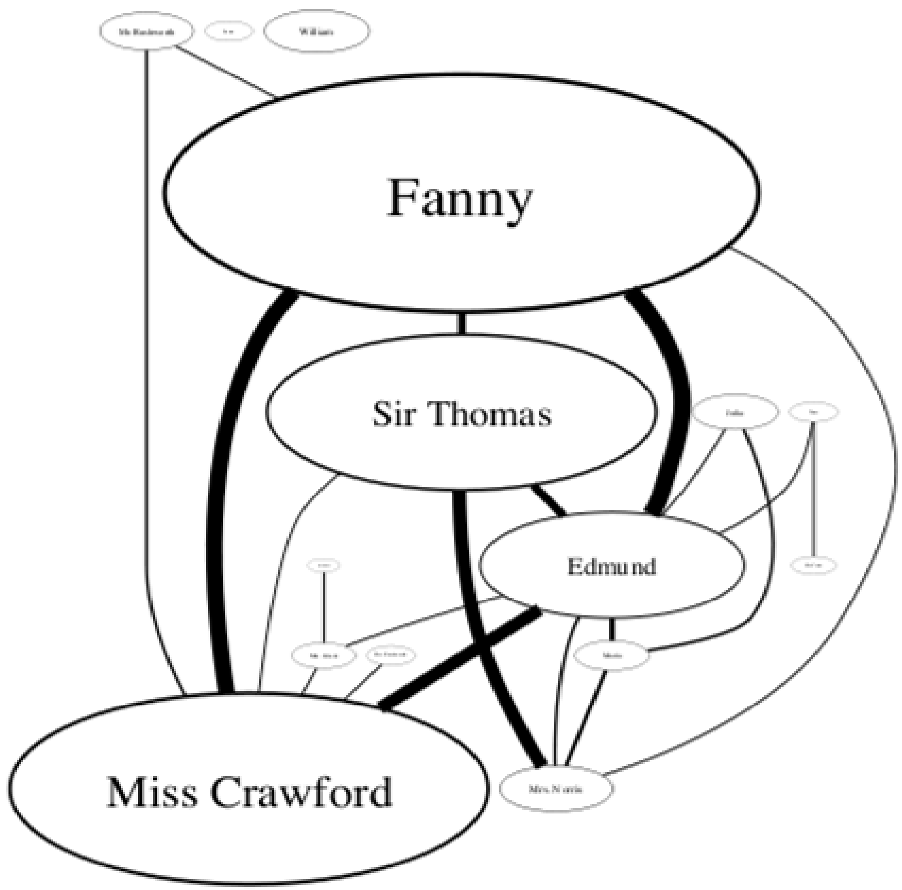


Figure 1: Automatically extracted conversation network for Jane Austen's *Mansfield Park*.

Elson, DK, Dames, N, & McKeown, KR (2010). Extracting social networks from literary fiction. In *Proceedings of the 48th annual meeting of the association for computational linguistics* (pp. 138-147). Association for Computational Linguistics.

Conversational networks

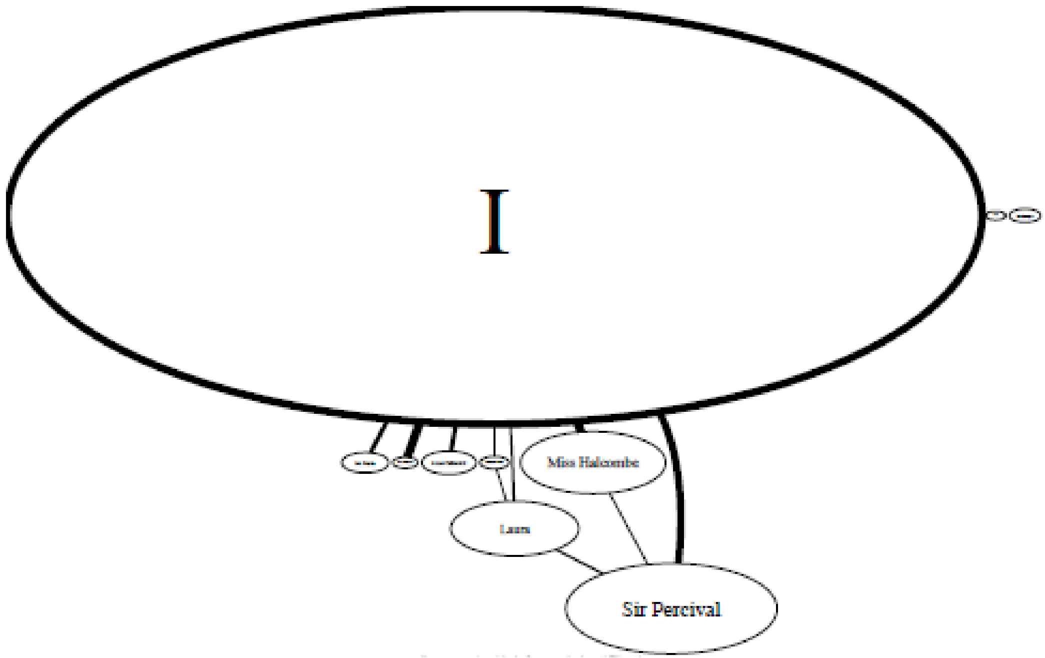


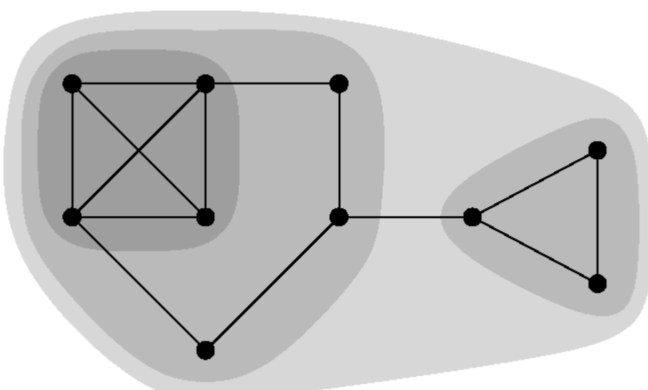
Figure 3: Conversational networks for first-person novels like Collins's *The Woman in White* are less connected due to the structure imposed by the perspective.

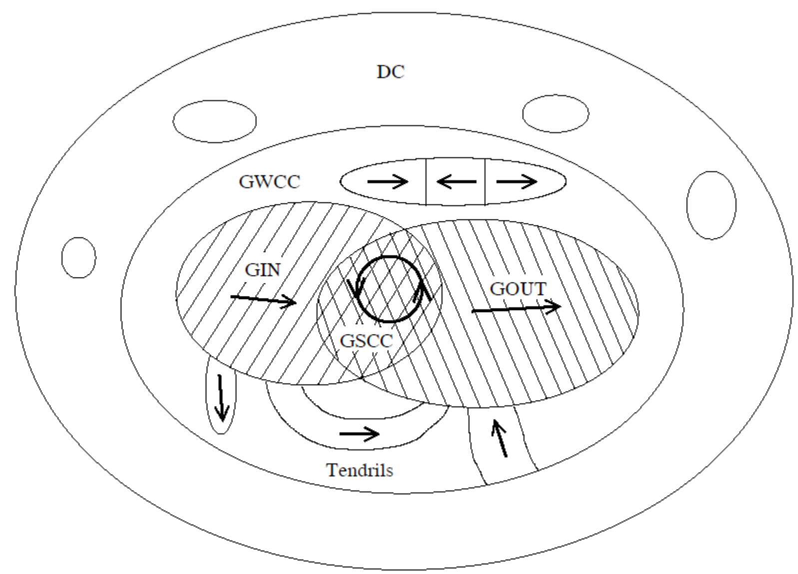
Elson, DK, Dames, N, & McKeown, KR (2010). Extracting social networks from literary fiction. In *Proceedings of the 48th annual meeting of the association for computational linguistics* (pp. 138-147). Association for Computational Linguistics.

NETWORK STRUCTURE AND MEASURES

Network structures

- (W/S)CC=(weakly/strongly) connected component
- core-periphery

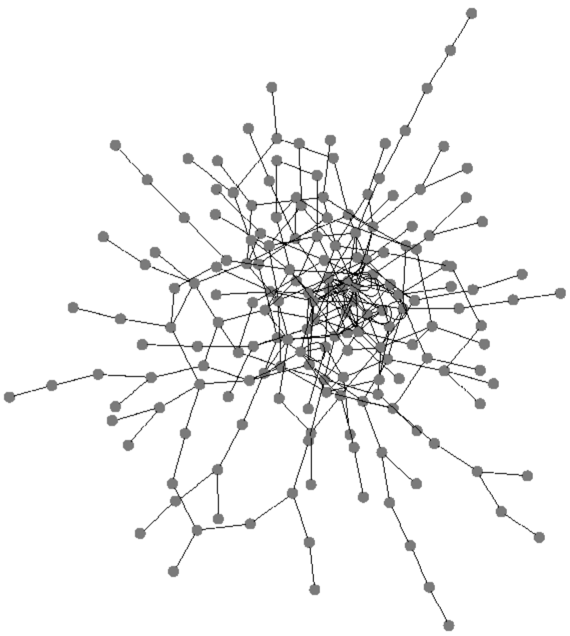




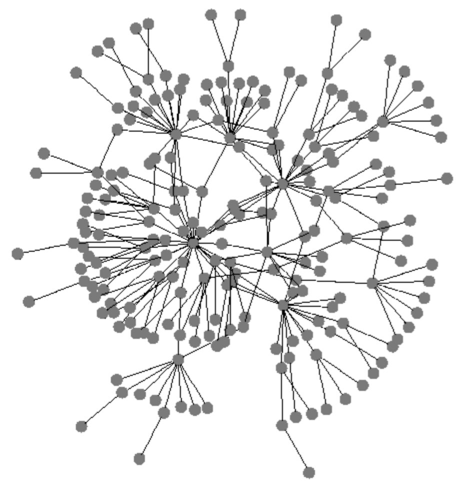
SN Dorogovtsev, JFF Mendes, *Evolution of networks,* Advances in Physics, **51**:4 (2002) 1079-1187
MEJ Newman, *Networks, An Introduction* (2010) Oxford University Press

Network structures

Assortative



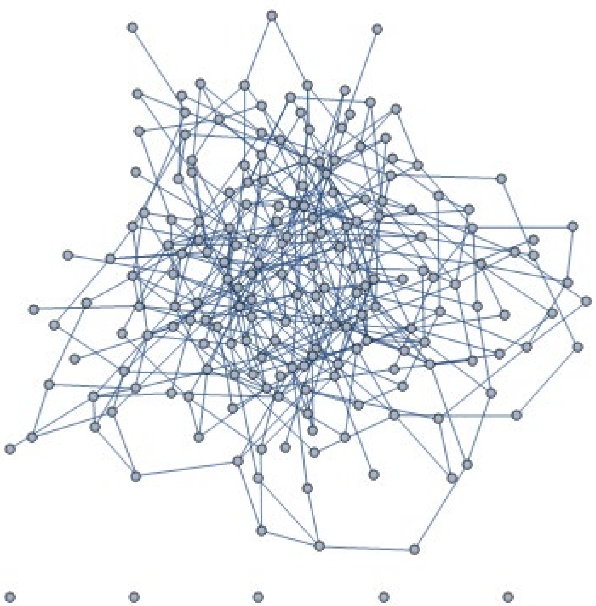
Disassortative



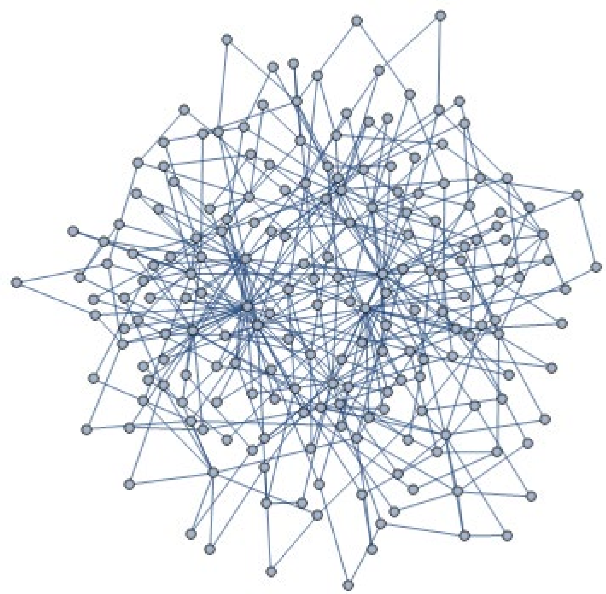
MEJ Newman, *Networks, An Introduction* (2010) Oxford University Press

Network models

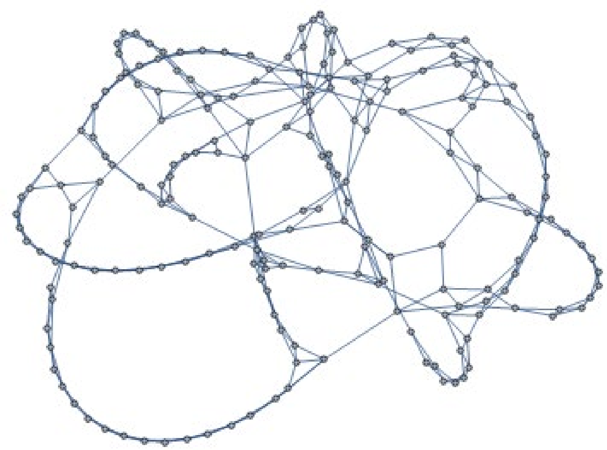
Random



Scale-free



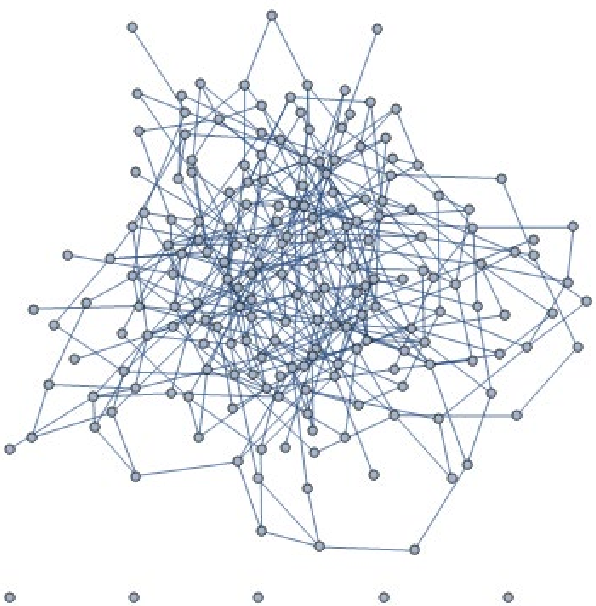
Small-world



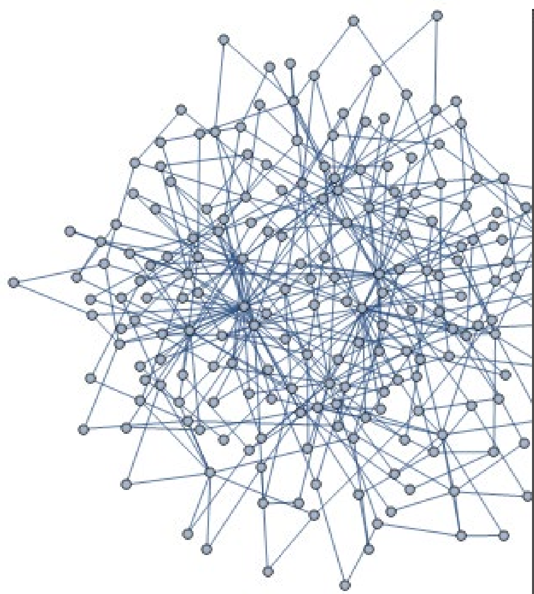
S. Milgram (1967) The small world problem, *Psychology Today* 2, pp. 60-67

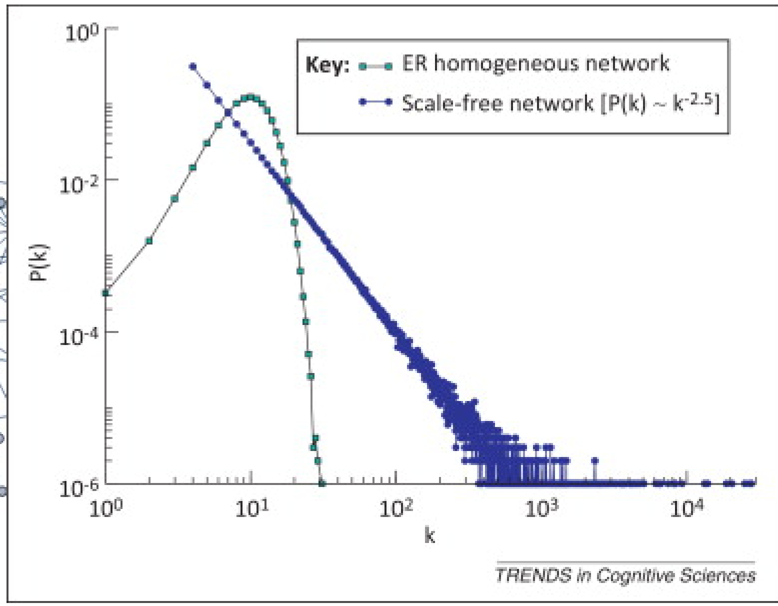
Network models

Random



Scale-free





Barabási AL, Albert R (1999). Emergence of scaling in random networks. *Science* 286 (5439), 509-512.
Baronchelli A, Ferrer-i-Cancho R, et al. (2013) Networks in Cognitive Science. Trends in Cognitive Sciences 17, pp. 348-360.

Centrality measures

How „important” is a node/link?

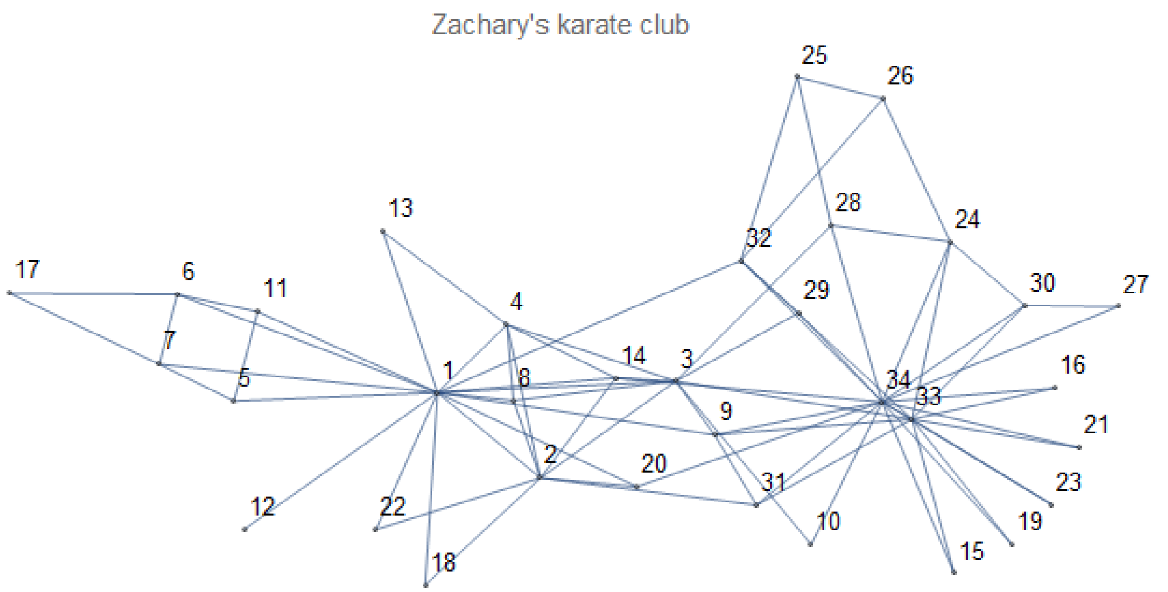
- degree
- clustering
- closeness
- betweenness
- …

MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Centrality measures

How „important” is a node/link?

- degree
- clustering
- closeness
- betweenness
- …

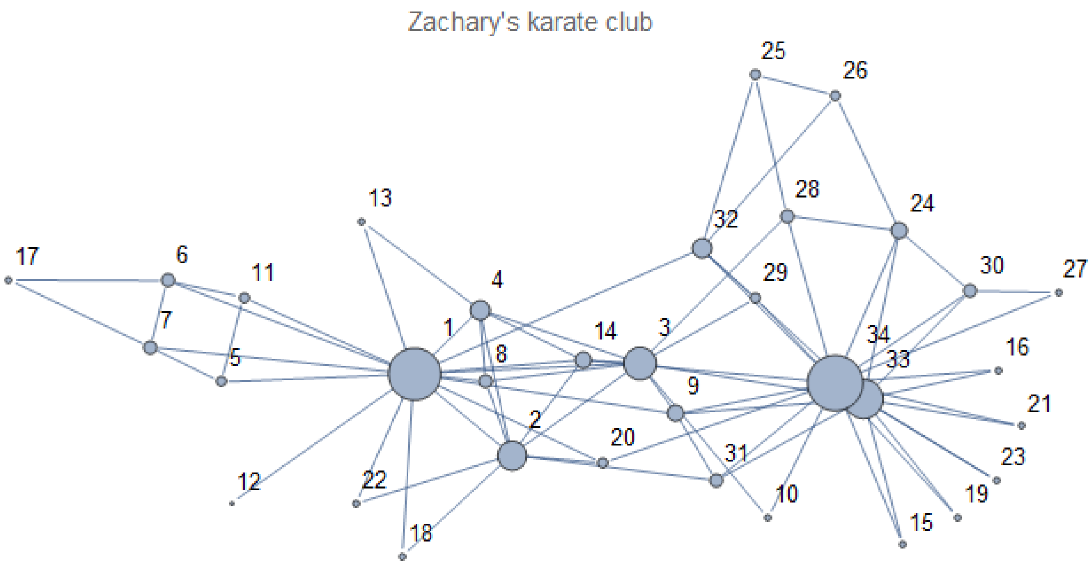


MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Centrality measures

How „important” is a node/link?

- **degree**
- clustering
- closeness
- betweenness
- …

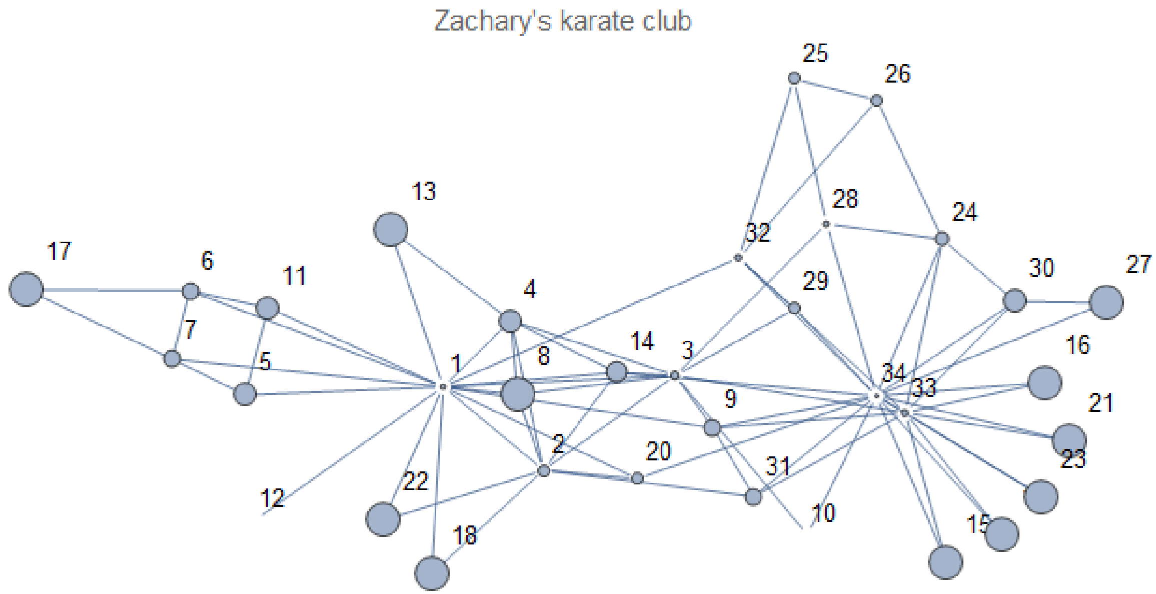


MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Centrality measures

How „important” is a node/link?

- degree
- **clustering**
- closeness
- betweenness
- …

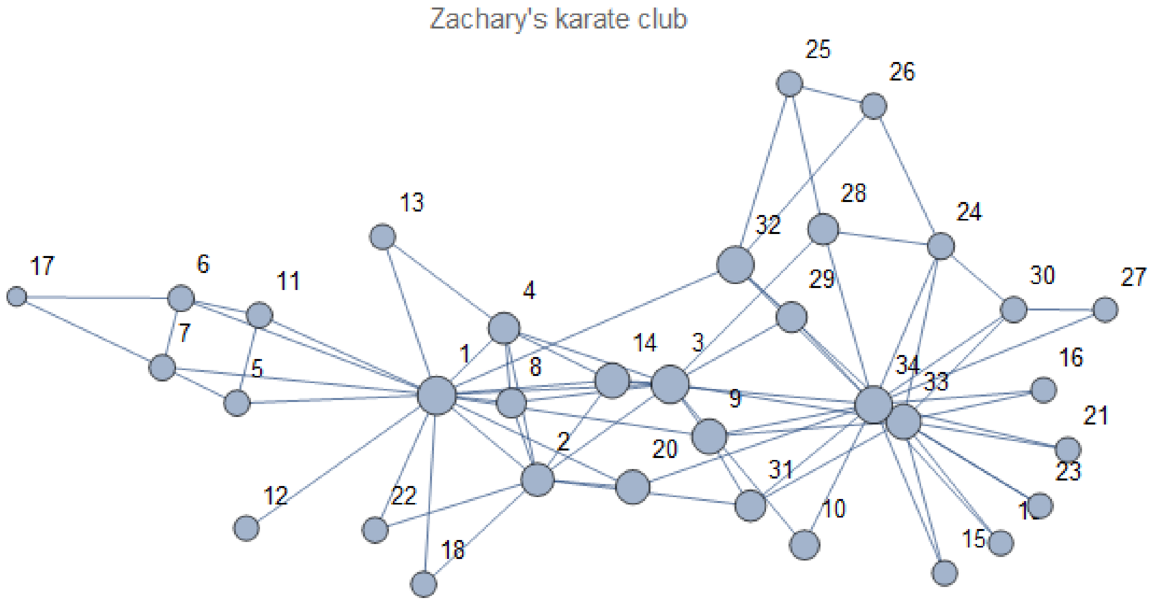


MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Centrality measures

How „important" is a node/link?

- degree
- clustering
- **closeness**
- betweenness
- …



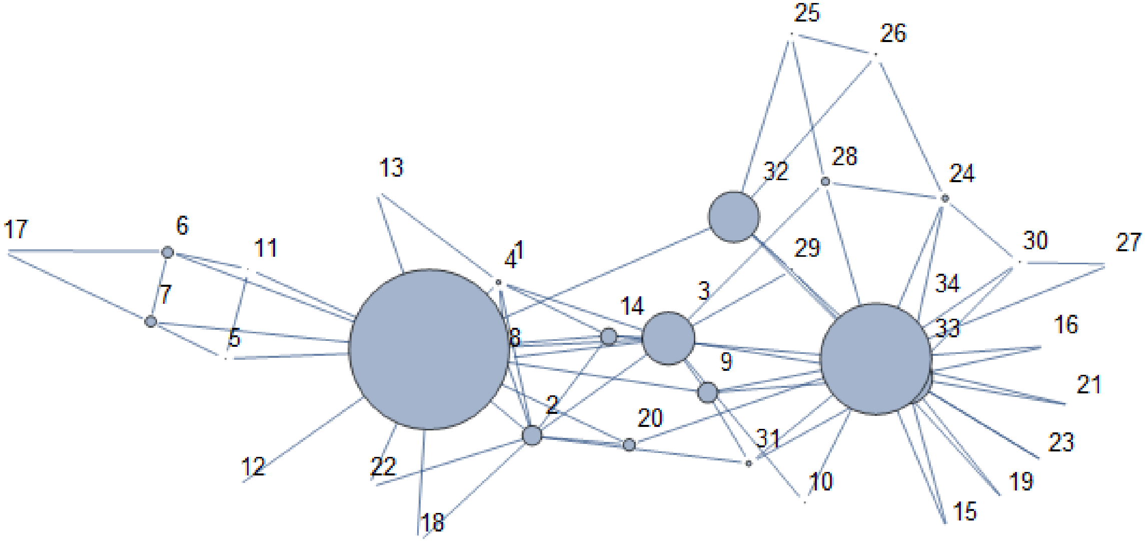
MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

Centrality measures

How „important” is a node/link?

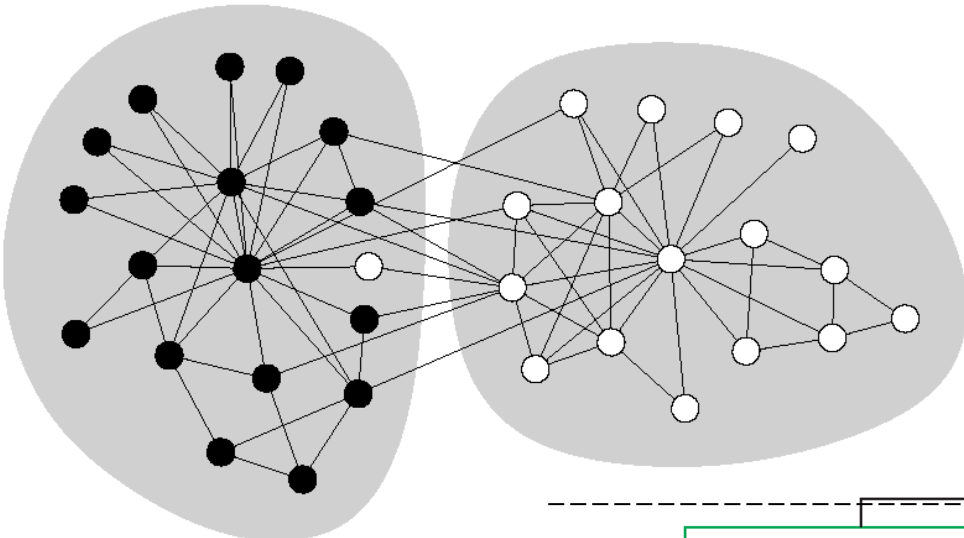
- degree
- clustering
- closeness
- **betweenness**
- …

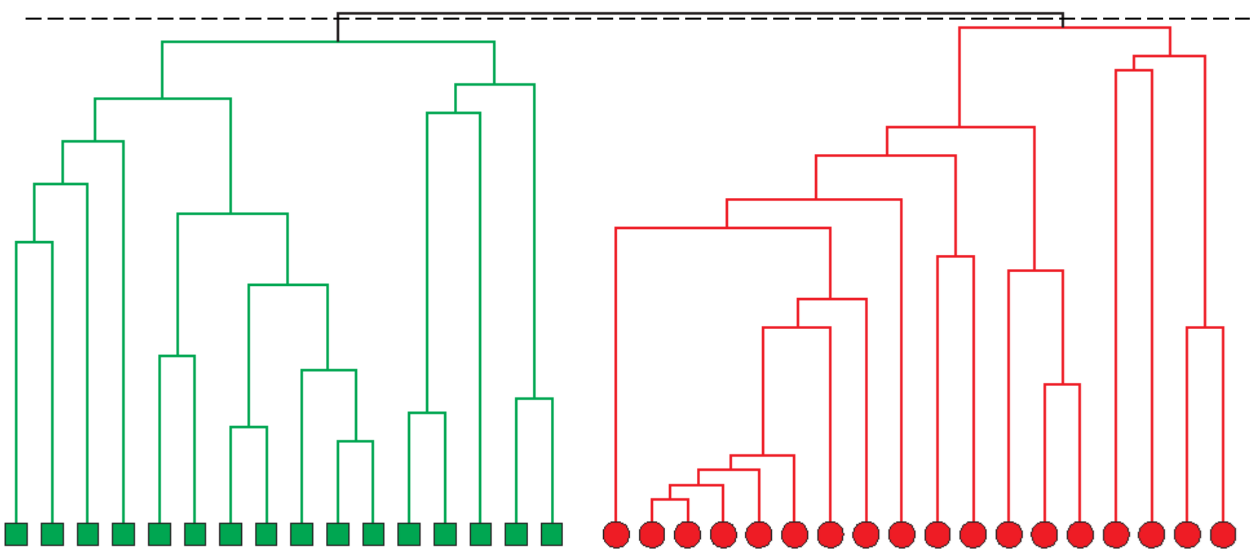
Zachary's karate club



MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256

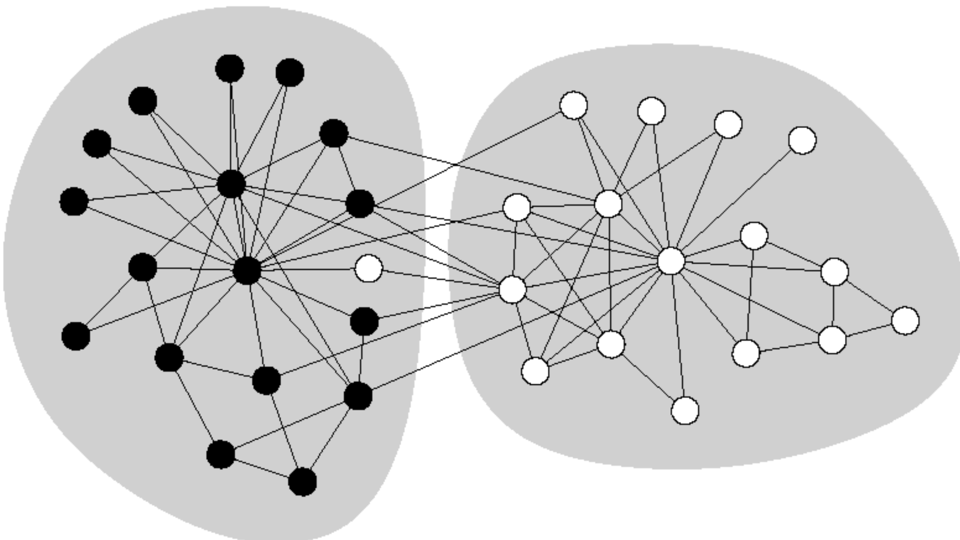
Community detection (clustering)





MEJ Newman, *The Structure and Function of Complex Networks,* SIAM REVIEW **45** (2003) 167–256
MEJ Newman, *Communities, modules and large-scale structure in networks*, Nature physics, 8 (2012) 25

Community detection (clustering)



$$Q = \frac{1}{2m}\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2m}\right)\delta(C_i, C_j)$$

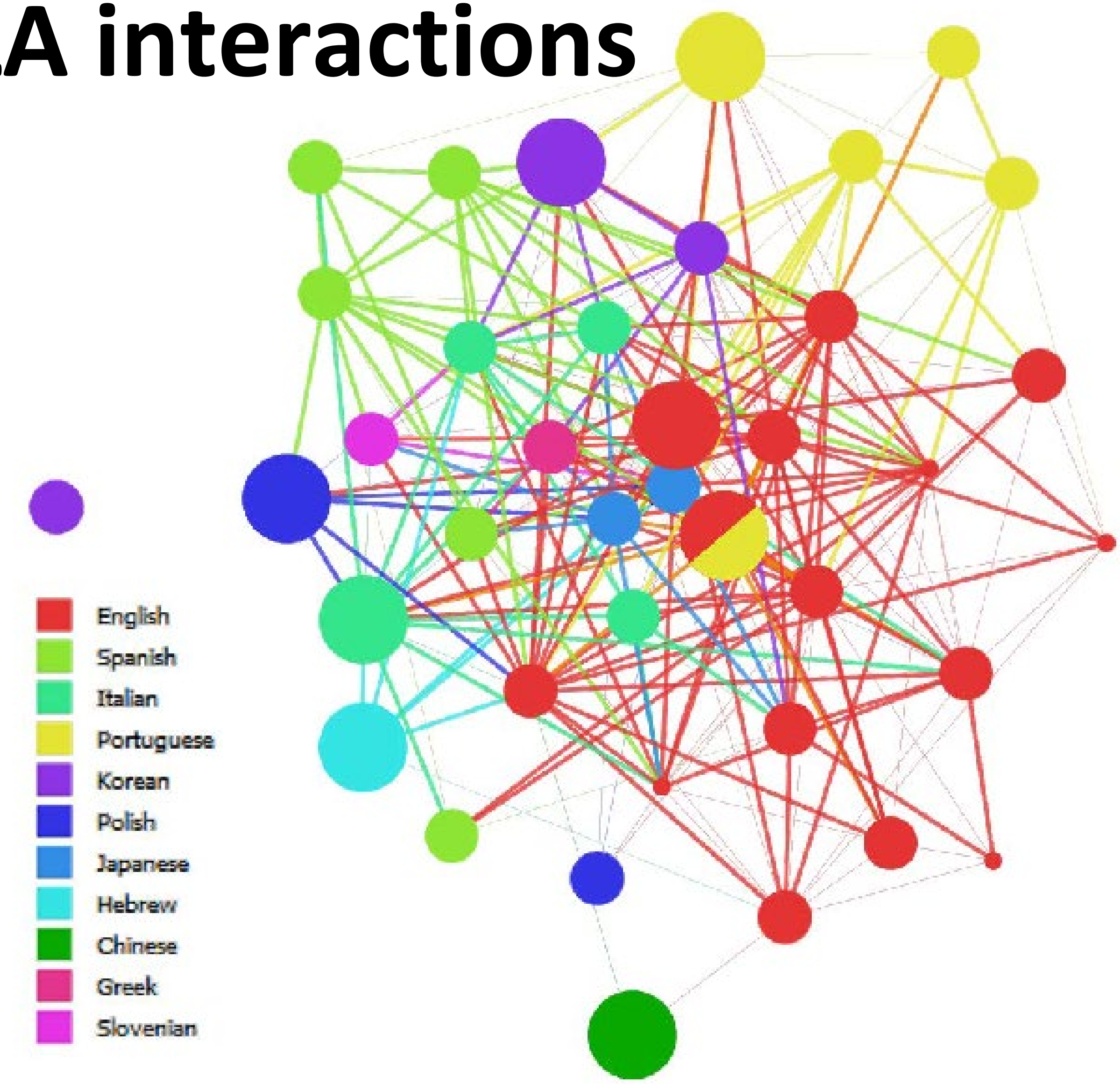
$$Q = \sum_{c=1}^{n_c}\left[\frac{l_c}{m} - \left(\frac{d_c}{2m}\right)^2\right]$$

	Nodes					
Nodes			1			
			1		1	
	1	1		1		1
			1		1	1
		1		1		
			1	1		

S Fortunato, *Community detection in graphs,* Physics Reports 486 (2010) 75-174

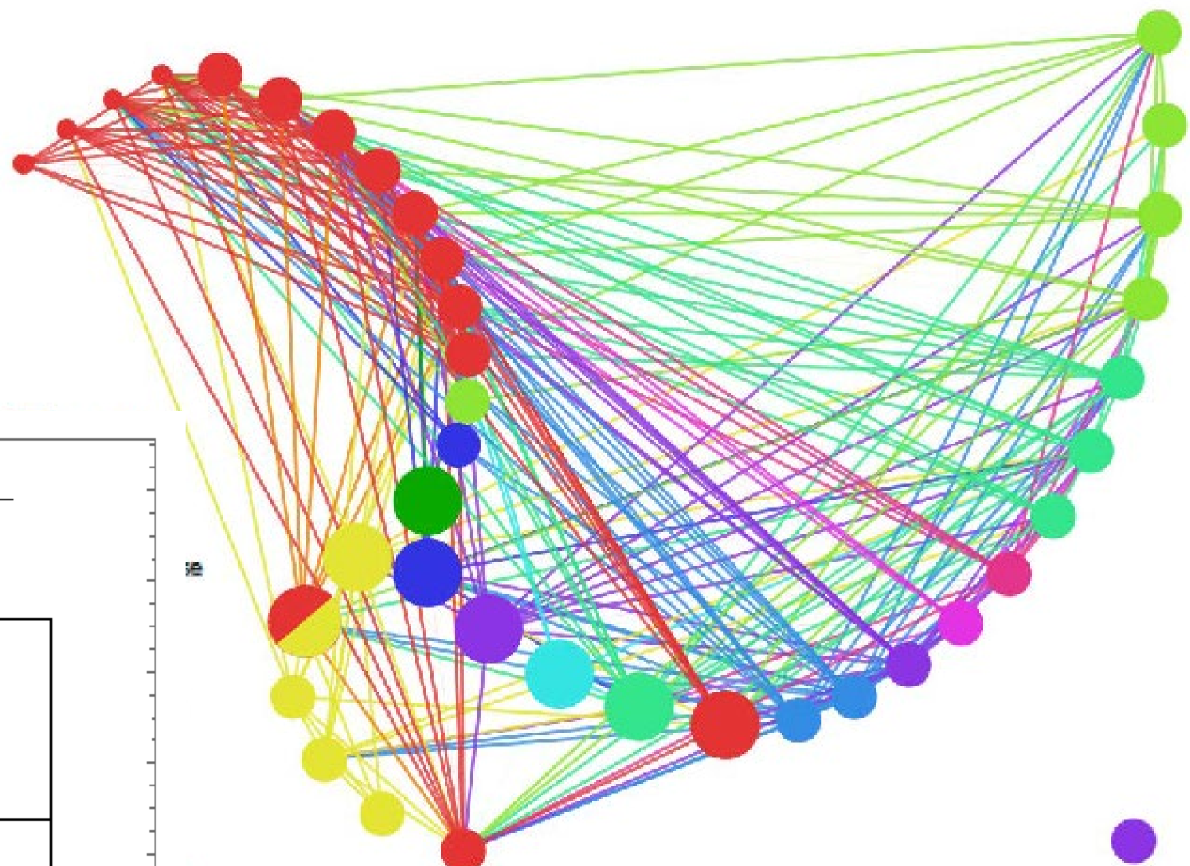
Example: SLA interactions

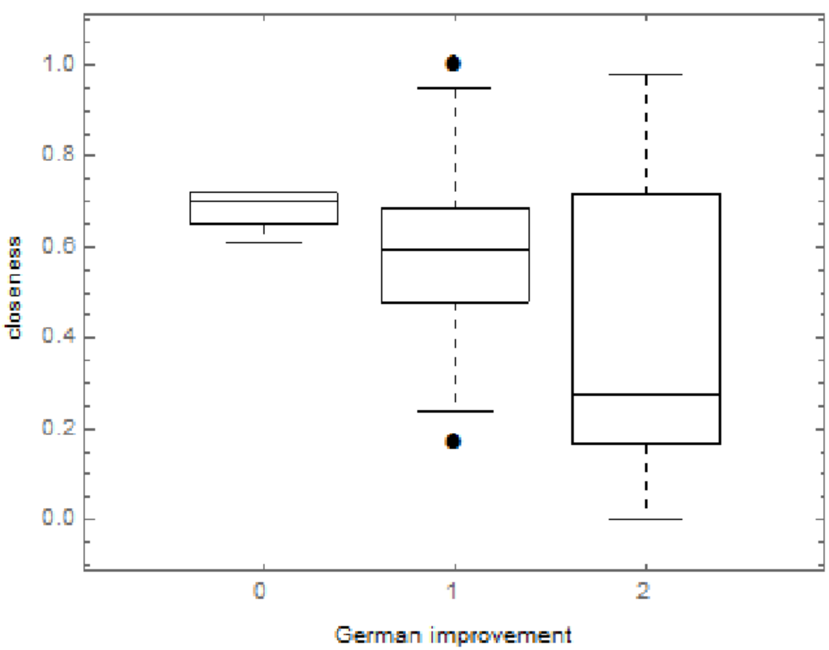
- branches= clusters
- node size= improvement

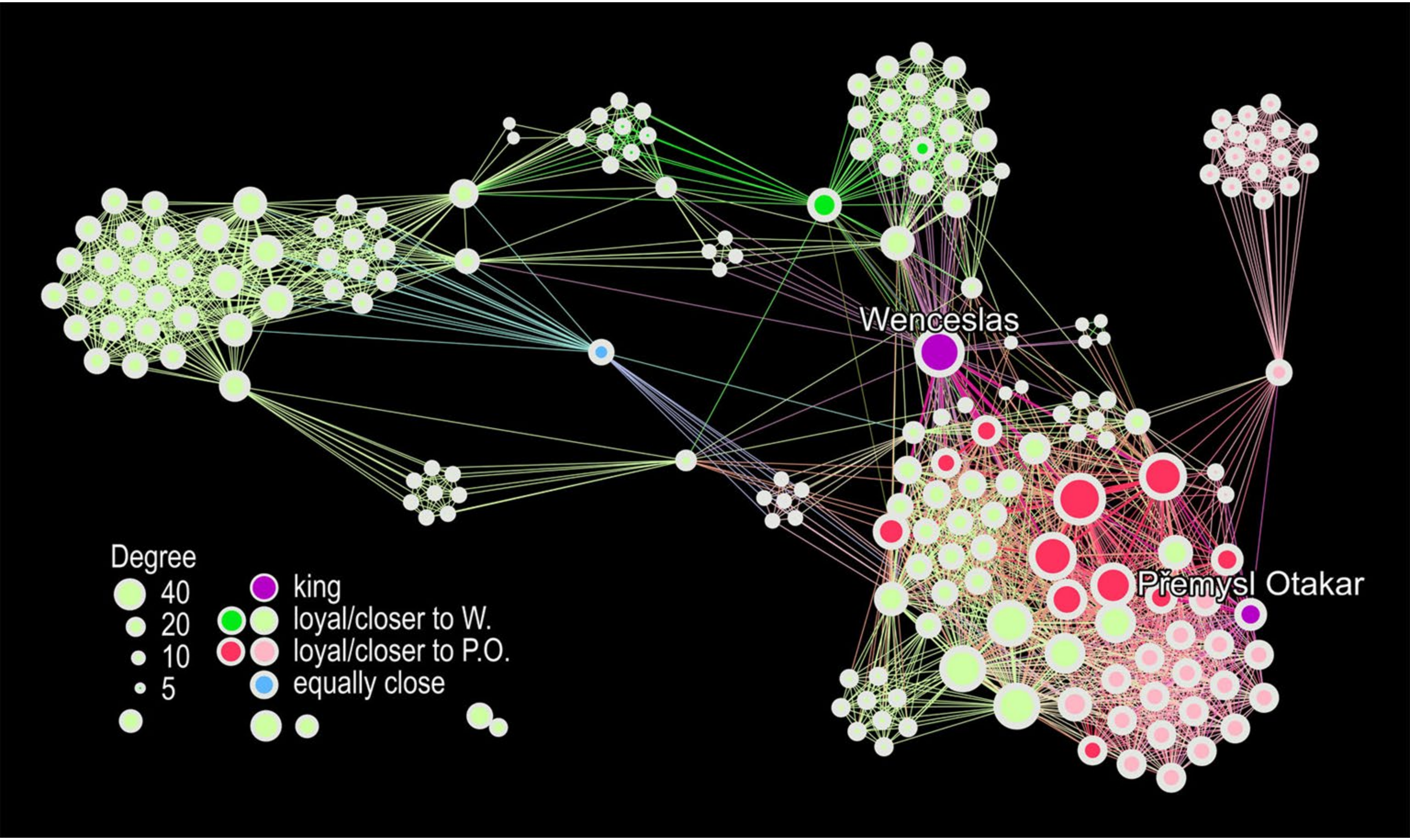


Example: SLA interactions

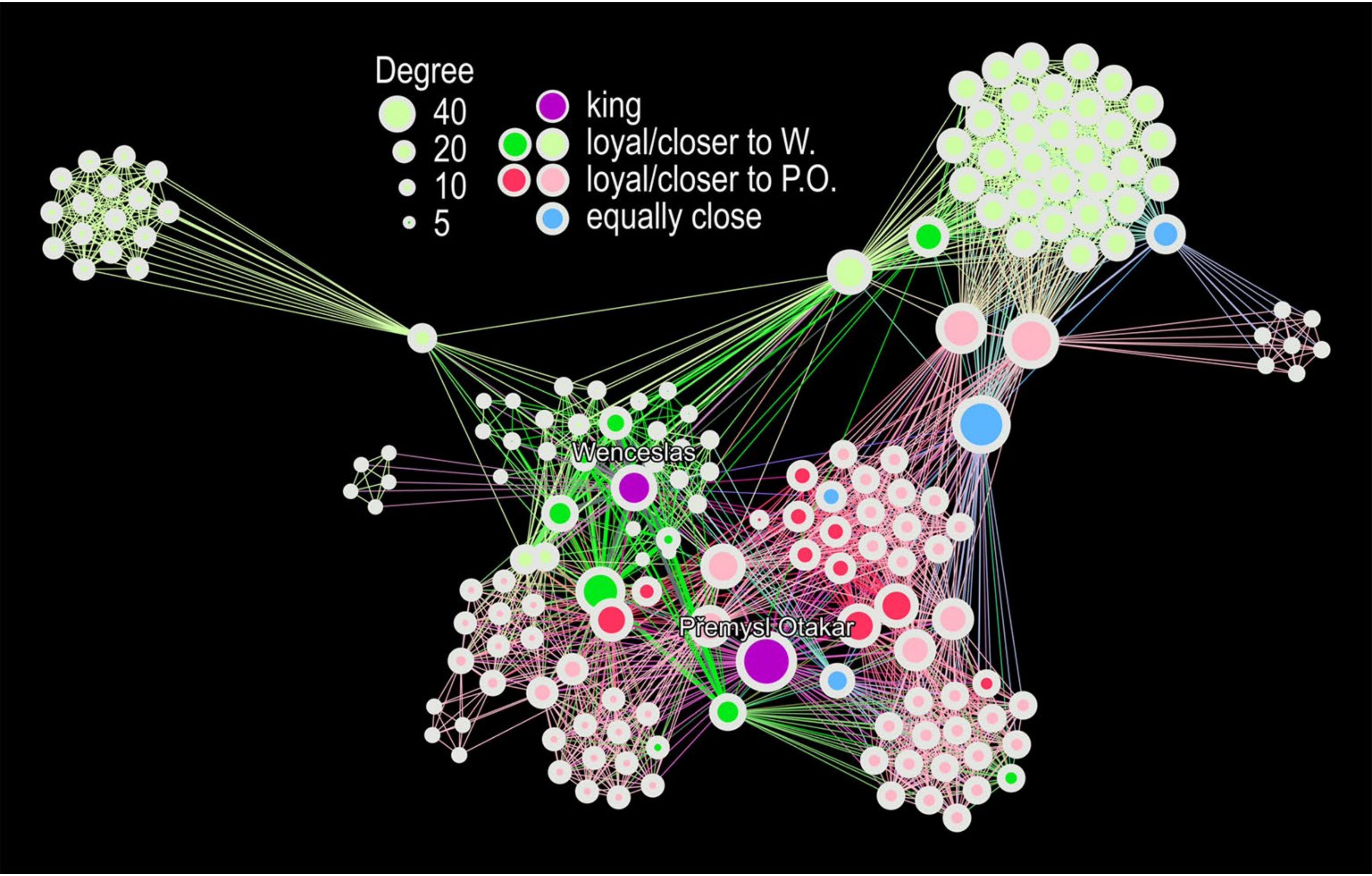
- branches= clusters
- node size= improvement

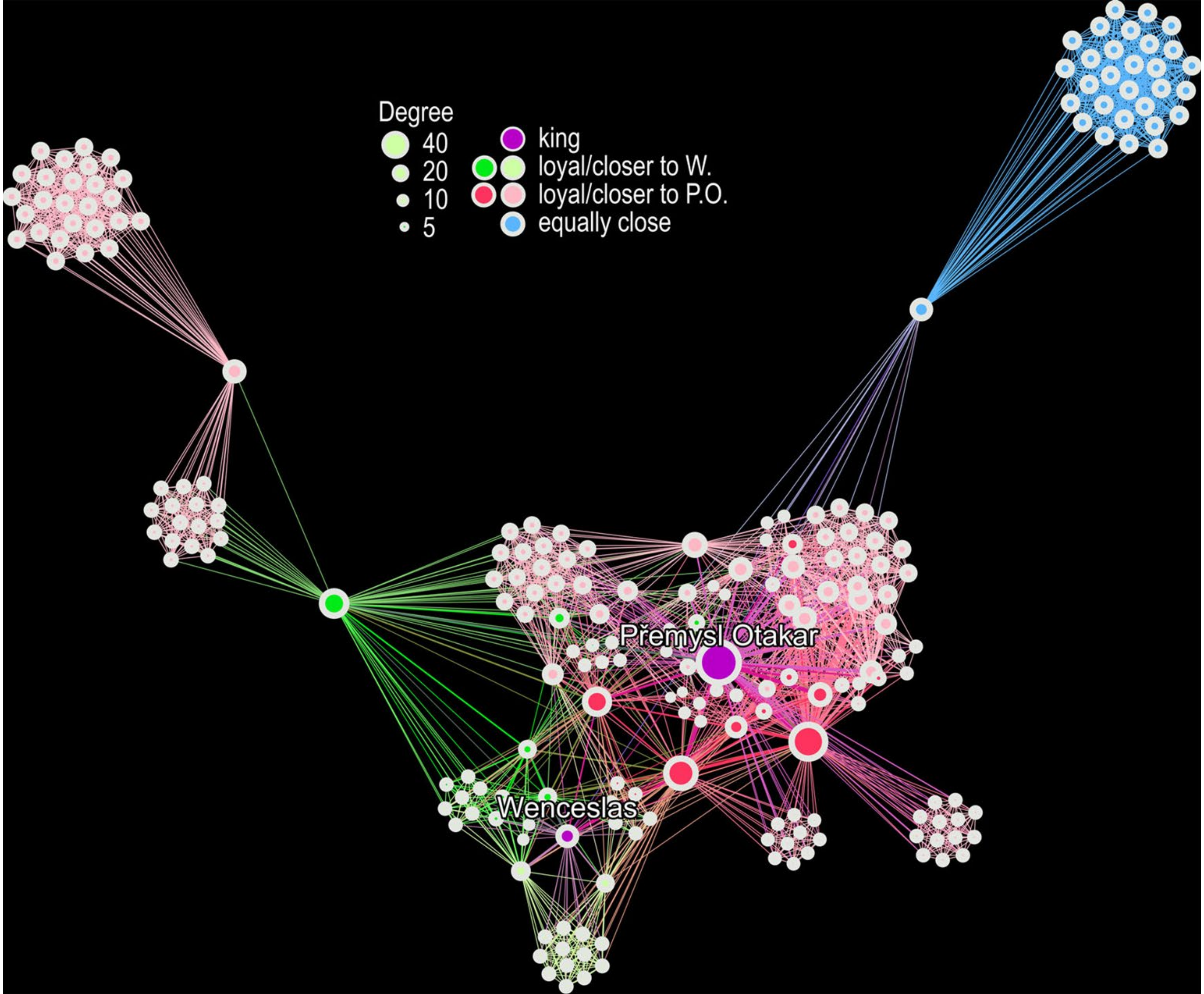


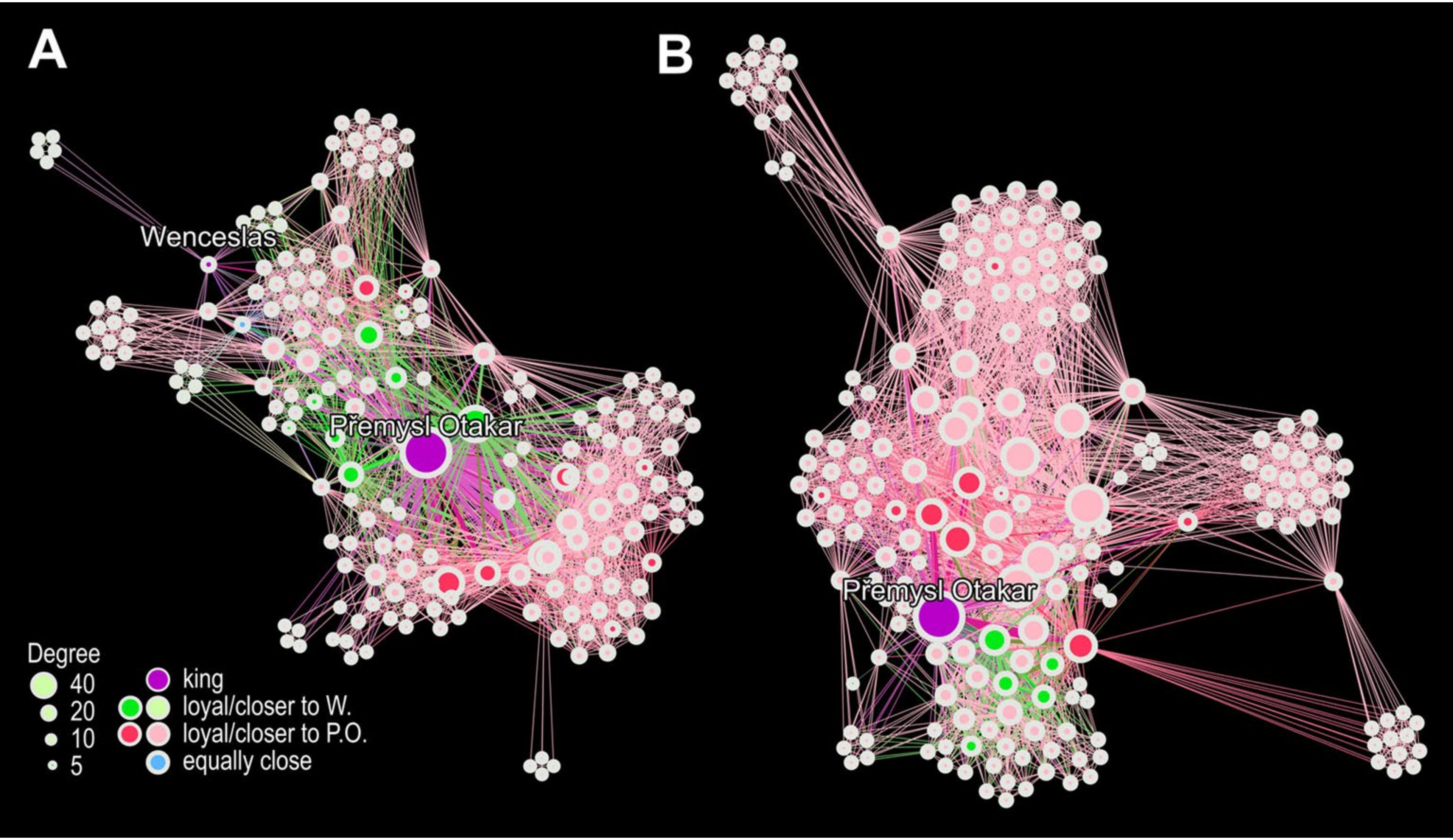


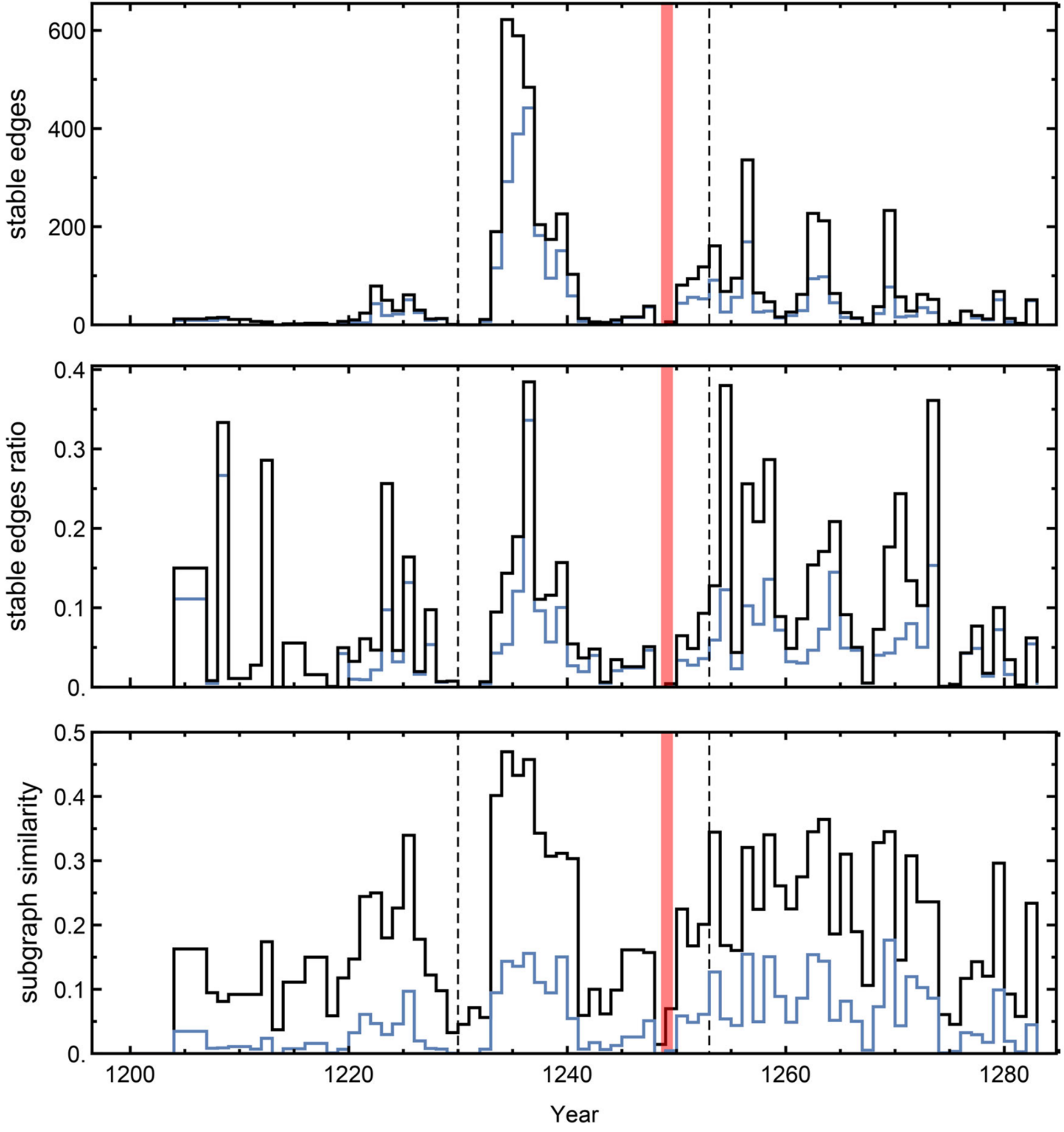


Wenceslas
Přemysl Otakar
Degree
40
20
10
5
king
loyal/closer to W.
loyal/closer to P.O.
equally close


Degree
40
20
10
5
king
loyal/closer to W.
loyal/closer to P.O.
equally close
Wenceslas
Přemysl Otakar


Degree
40
20
10
5
king
loyal/closer to W.
loyal/closer to P.O.
equally close
Přemysl Otakar
Wenceslas


A
B
Wenceslas
Přemysl Otakar
Přemysl Otakar
Degree
40
20
10
5
king
loyal/closer to W.
loyal/closer to P.O.
equally close



stable edges
600
400
200
0
stable edges ratio
0.4
0.3
0.2
0.1
0.
subgraph similarity
0.5
0.4
0.3
0.2
0.1
0.
1200
1220
1240
1260
1280
Year