Hierartical Structure in Social Networks

Cindy Cook, Matt Denny, Mitch Goist, Timmy Huynh

SODA502: Final Project

9 December 2015

Theorizing Hierarchy in Networks - Prior Work

Metaphor/Measure	Relevant Authors	Visual Representation
1. Web of group affiliation, groups versus networks	Simmel [1922] 1955	
2. Sociograms, sociometric stars	Moreno 1934	
3. Structural equivalence	White, Boorman, and Breiger 1974; Burt 1992	
4. Strength of ties, weak versus strong	Granovetter 1973	
5. Bridges, structural holes, *tertius gaudens*	Burt 1992	
6. Degrees of separation, path length	Milgram 1967	
7. Interlock centrality, sphere of influence	Mizruchi 1996	Third Board of Directors; Competitor Board 1; Competitor Board 2
8. Local versus global network, closeness centrality	Freeman 1979	
9. Exchange, centrality versus power (A is central, but B has power because both C and D depend on B)	Bonacich 1987; Cook 1977	A; B; C; D
10. Density	Barnes 1979; Marsden 1993	
11. Small-worlds	Watts and Strogatz 1998, 441	Regular; Small-World; Random; $p = 0$; Increasing Randomness; $p = 1$

FIGURE 1. Social network conceptual toolkit. Note: Figure courtesy of *Nature*, vol. 363.

Theorizing Hierarchy in Networks - Mann

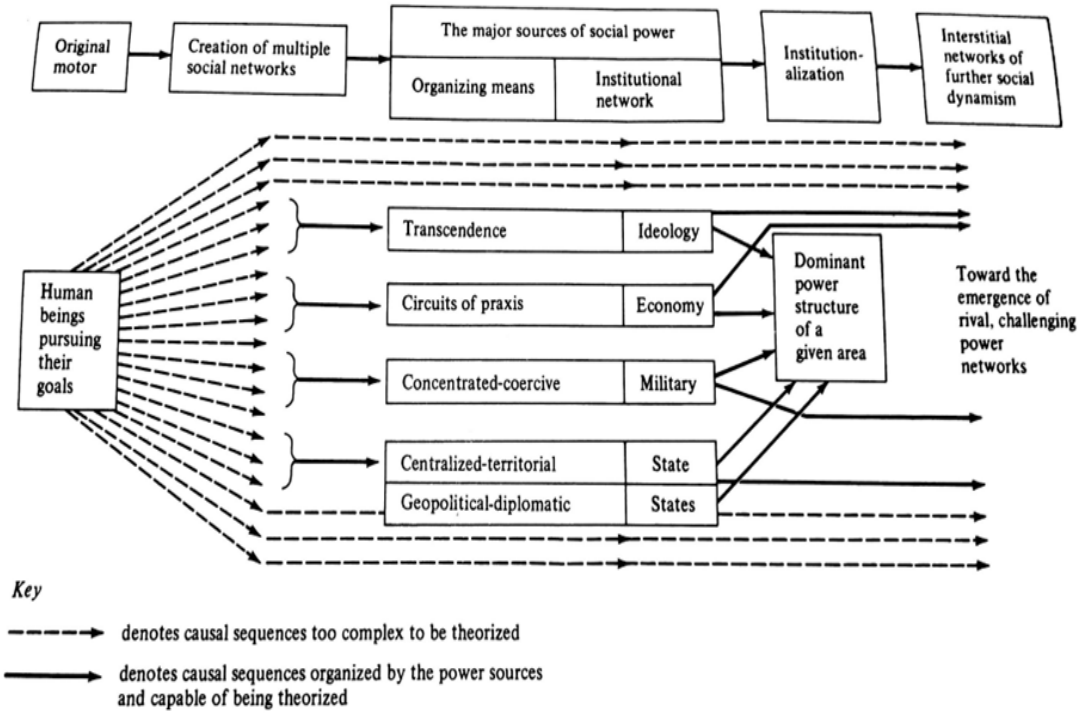


Figure 1.2. *Causal IEMP model of organized power*

Societies as organized power networks

	Authoritative	*Diffused*
Intensive	Army command structure	A general strike
Extensive	Militaristic empire	Market exchang

Figure 1.1. *Forms of organizational reach*

Hierarchy in networks–Measures

Measure	Undirected	Weighted	Global	Local
Landau's h		X	X	
Kendall's K		X	X	
Reach Degree	X	X	X	X
Reach Closeness	X	X	X	X
GRC	X	X	X	X
Rooted Depth			X	
Degree	X	X	X	X
Closeness	X	X	X	X
Betweenness	X	X	X	X
Eigenvector	X	X	X	X

Hierarchy in Networks–Simulated Datasets

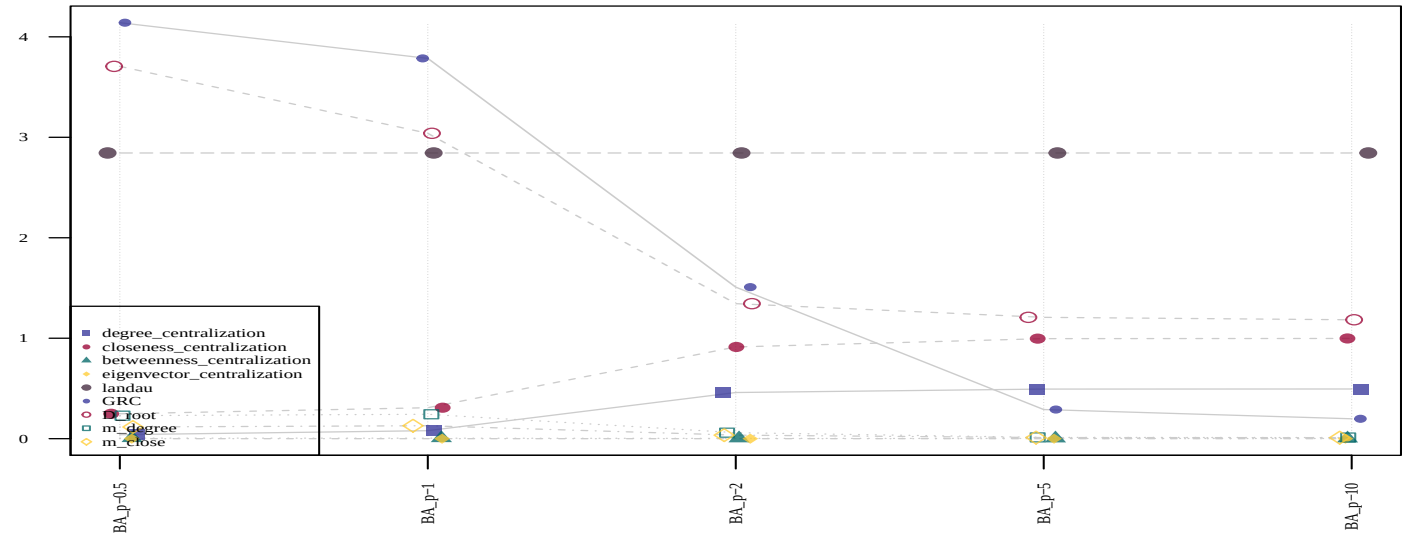
- 7500 Barabasi-Albert (BA) Datasets with node size $50, 200$, or 500 and preferential attachment parameter $0.5, 1, 2, 5$, or 10.
- 6000 Tree-Structured (TR) Datasets with node size $50, 200$, or 500 and children parameter $2, 5, 10$, or 50.
- 4500 Erdos-Renyi (ER) Datasets with node size $50, 200$, or 500 and parameter size $0.05, 0.1$, or 0.2.

BA and TR	Node	Edge	Density
	50	49	0.02
	200	199	0.005
	500	499	0.002

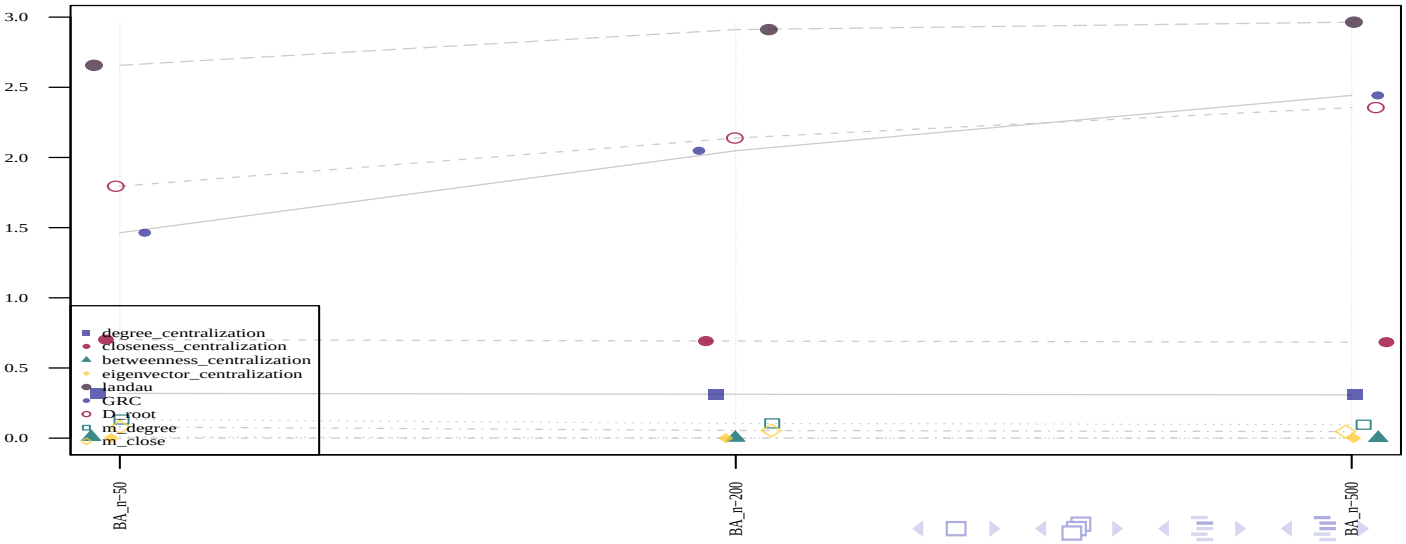
ER	Node	Parameter	Edge	Density	Cluster Coefficient
	50	0.05	122.524	0.050	0.0956
	50	0.1	244.768	0.100	0.188
	50	0.2	490.050	0.200	0.359
	200	0.05	1989.242	0.050	0.098
	200	0.1	3980.776	0.100	0.190
	200	0.2	7952.138	0.120	0.360
	500	0.05	12473.270	0.050	0.097
	500	0.1	24932.824	0.100	0.190
	500	0.2	49892.784	0.200	0.360

Average Global Hierarchy–BA Networks

Parameters

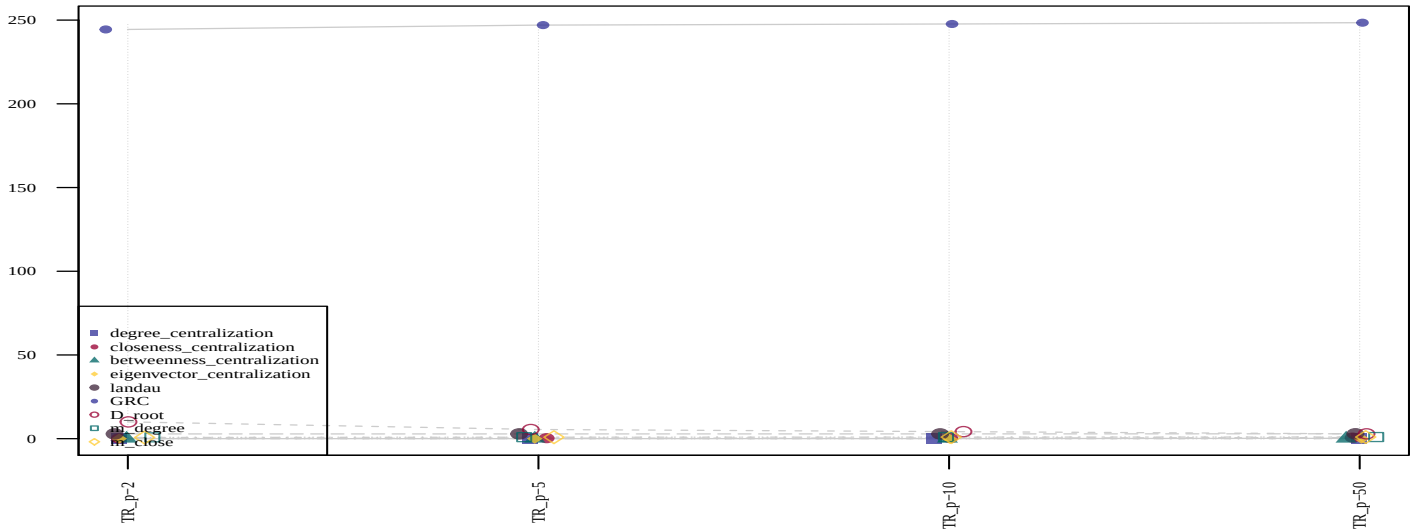


Node Size

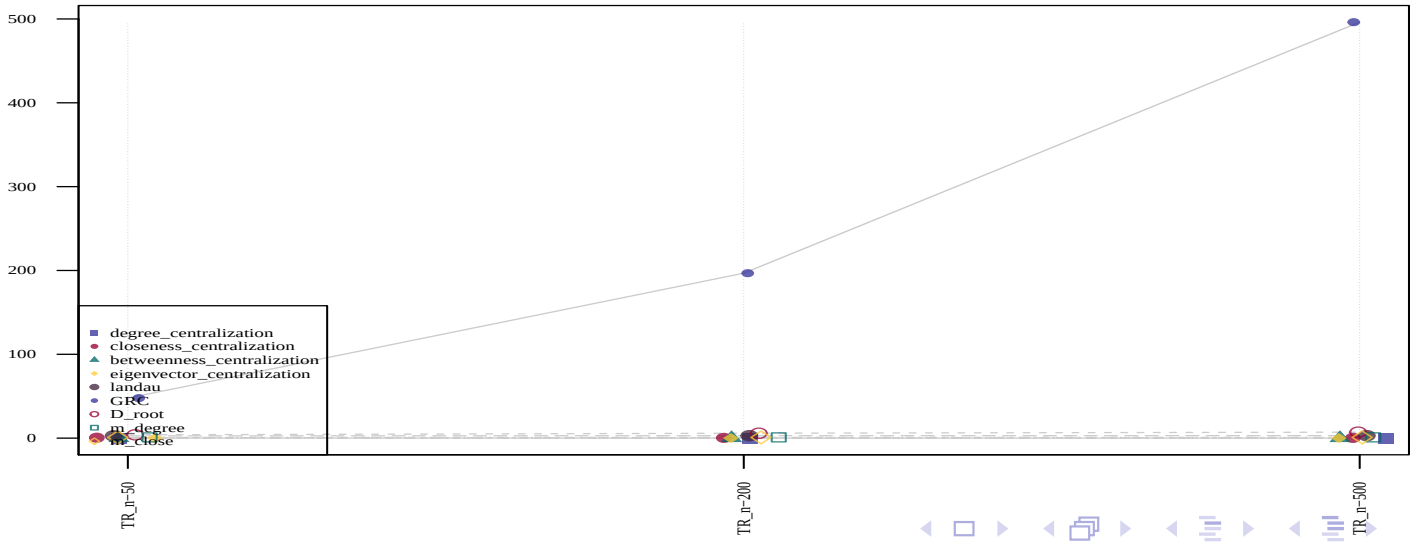


Average Global Hierarchy–TR Networks

Parameters

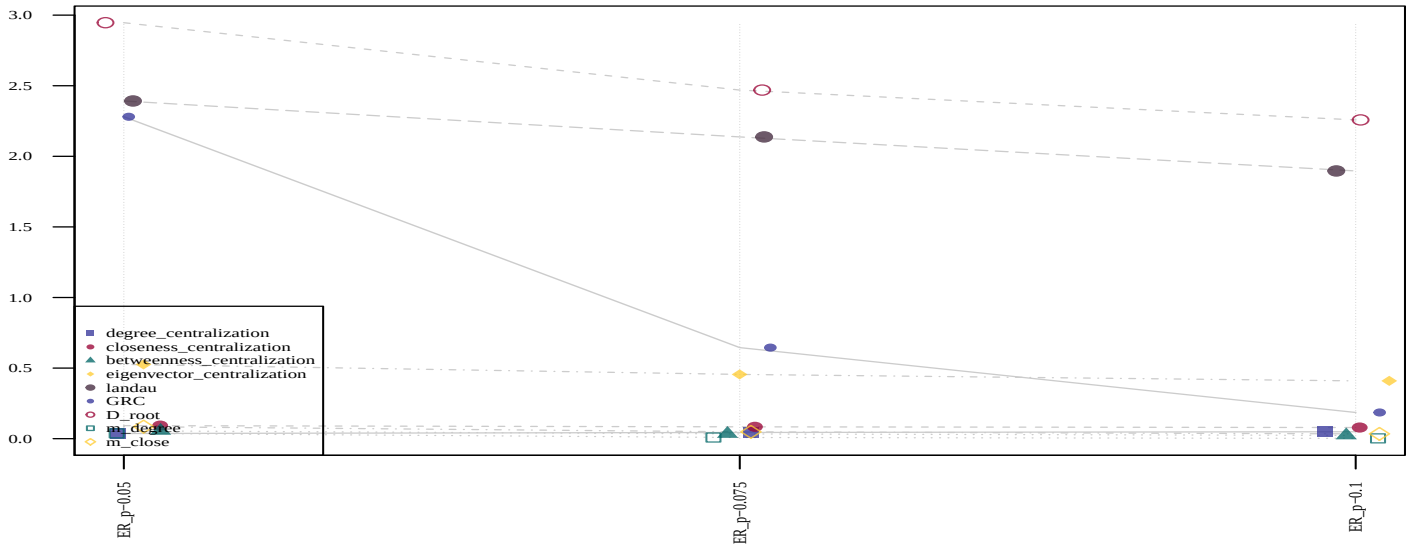


Node Size

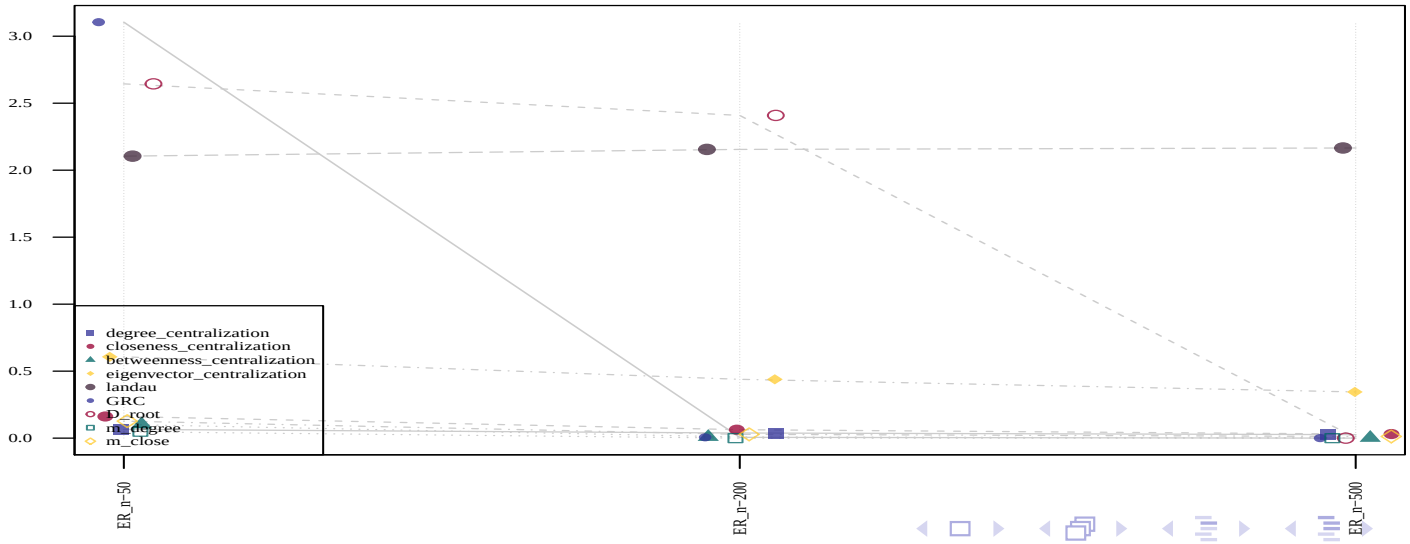


Average Global Hierarchy–ER Networks

Parameters



Node Size



Hierarchy in Networks–Results

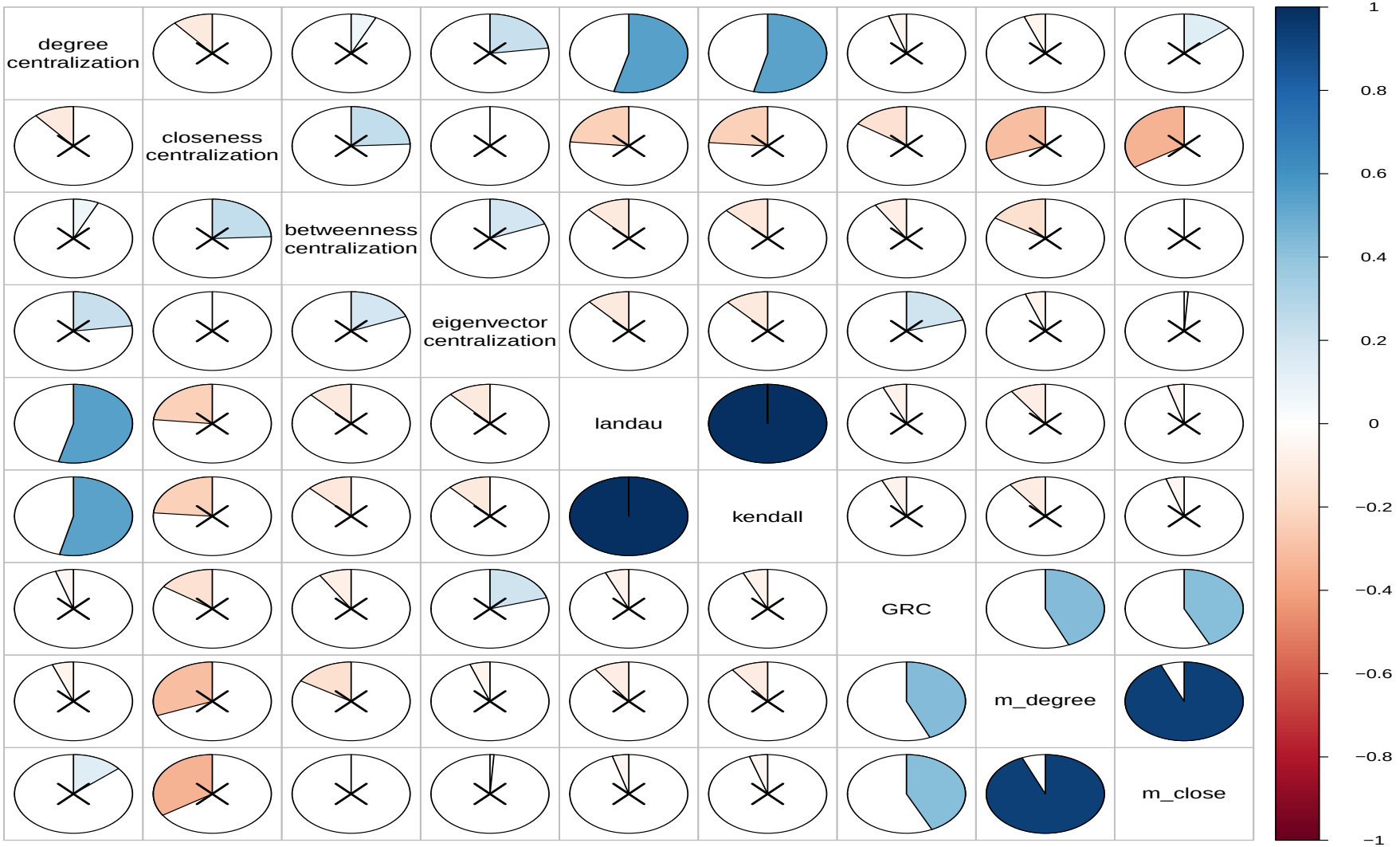
- GRC and D Root are not robust to changes in node size.
  - GRC increases with node size for both BA and TR networks
  - GRC decreases with node size for ER networks.
  - D Root increases with node size for BA networks.

Hierarchy in Networks–Datasets

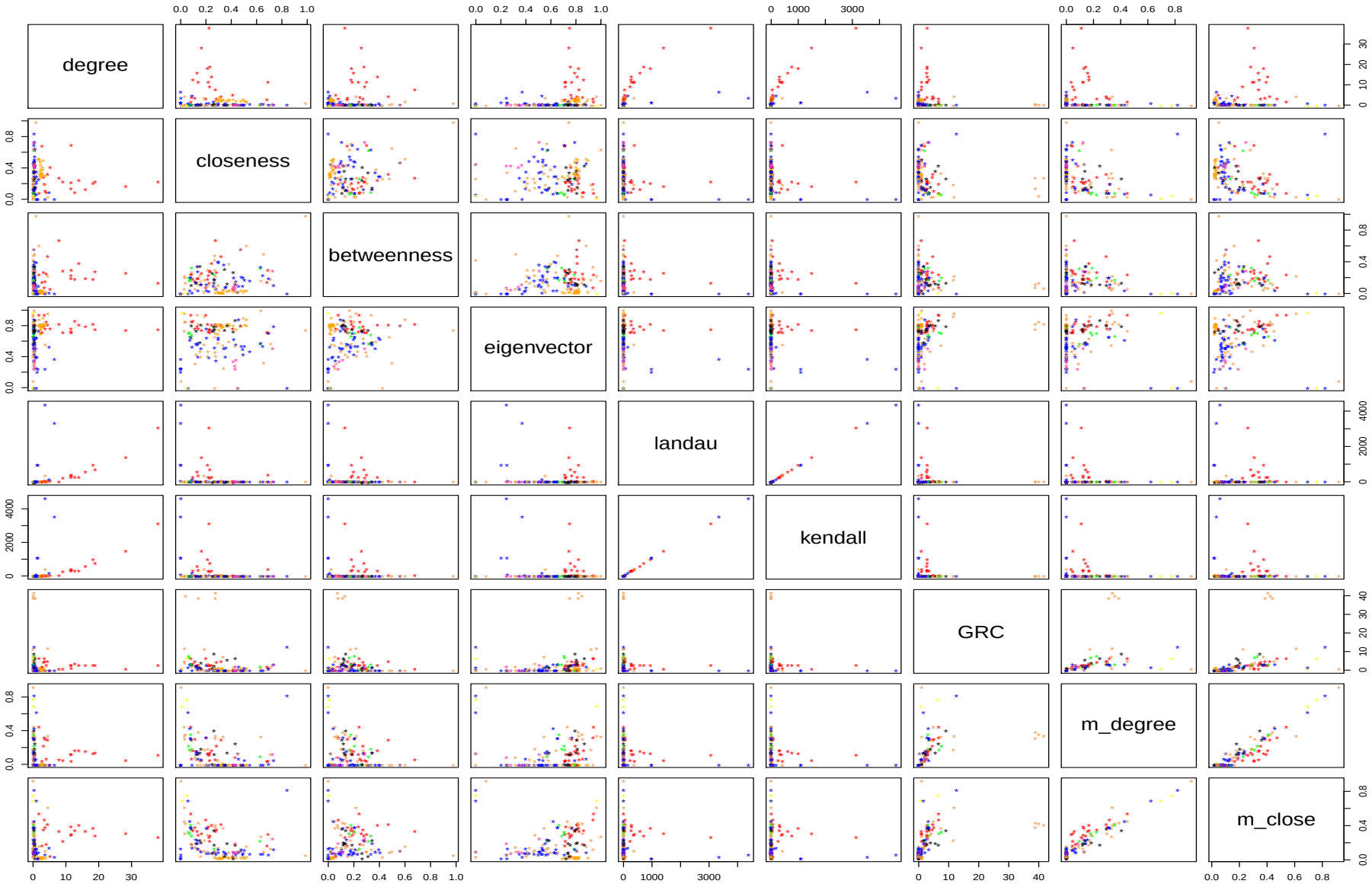
- 141 datasets from UCINET, organizational emails, and Congress

Type	Nodes	Edges	Density	Cluster Coeff.
Communication	25.714	2419	3.483	0.558
Cosponsorship	101.222	13358.889	1.317	0.788
Co-membership	22	9267.333	18.318	0.377
Interaction	23.075	1944.95	1.985	0.589
Unknown	76.833	4688.25	0.360	NaN
Friendship	29.833	92	0.122	0.348
Affect	17.636	95.182	0.319	0.329
Terrorism	63	308	0.079	0.361
Trade	24	285.6	0.5170	0.734

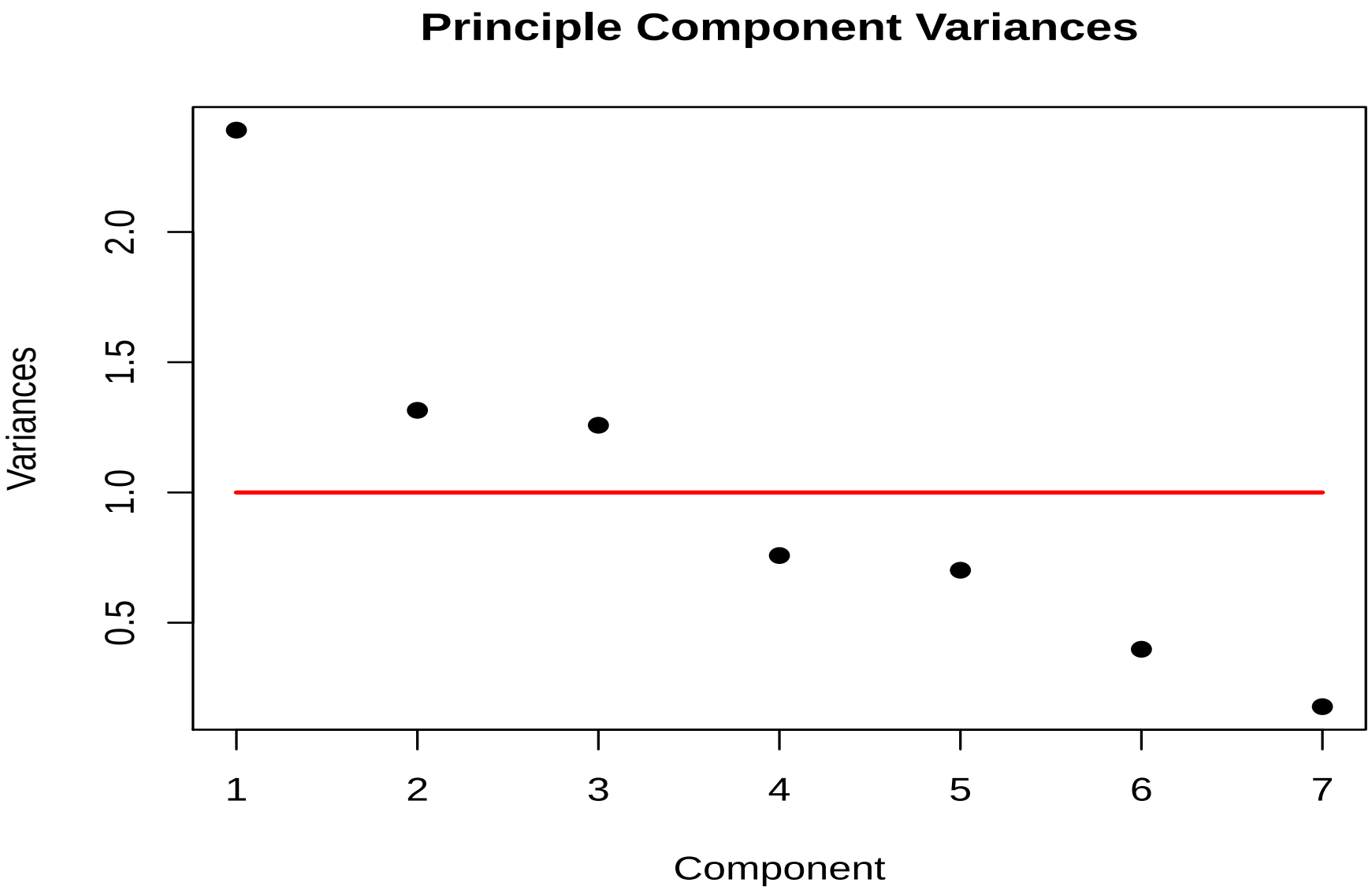
Hierarchy in Networks – Global Measures



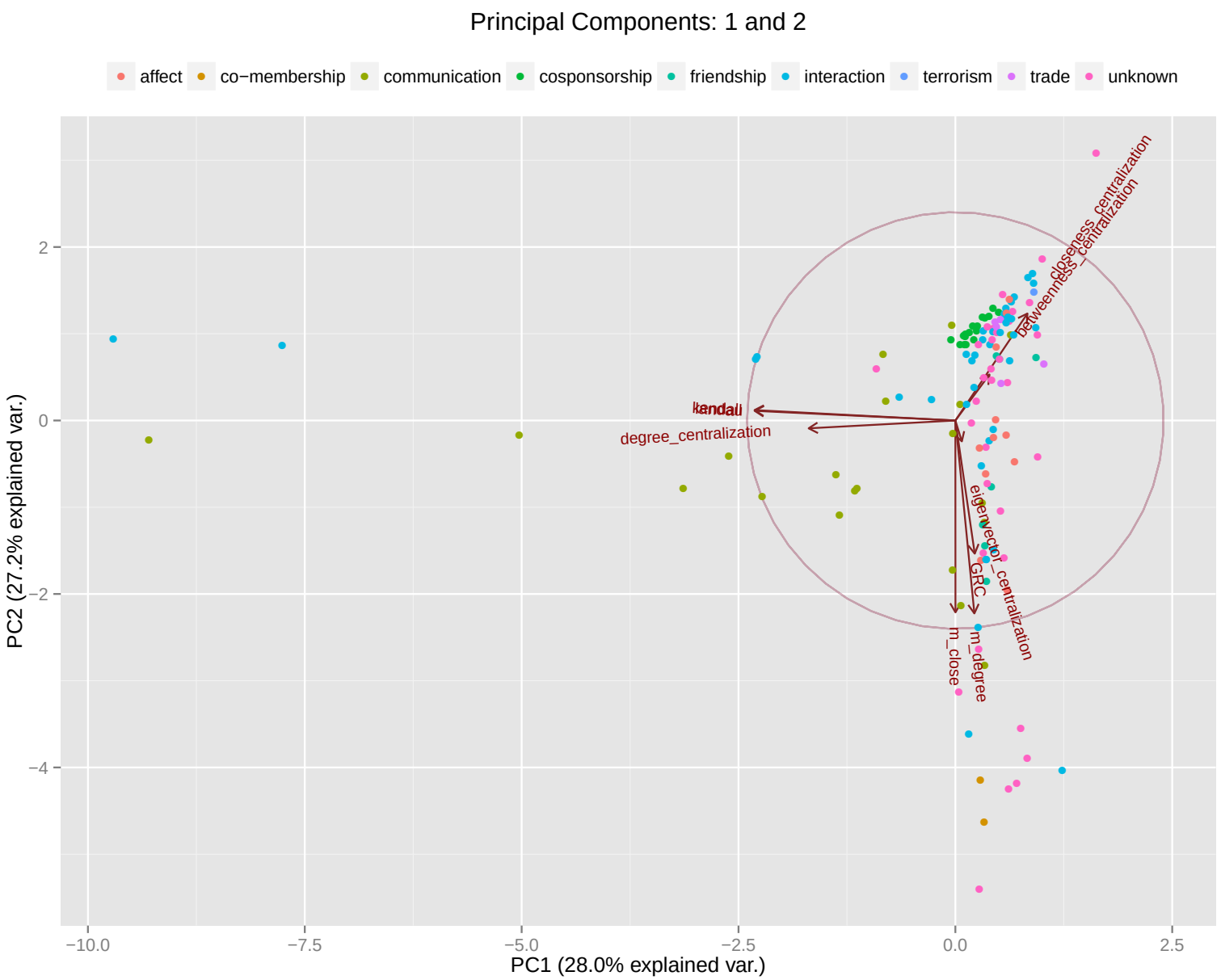
Hierarchy in Networks – Global Measures



Hiearchy in Networks – PCA



Hierarchy in Networks – PCA



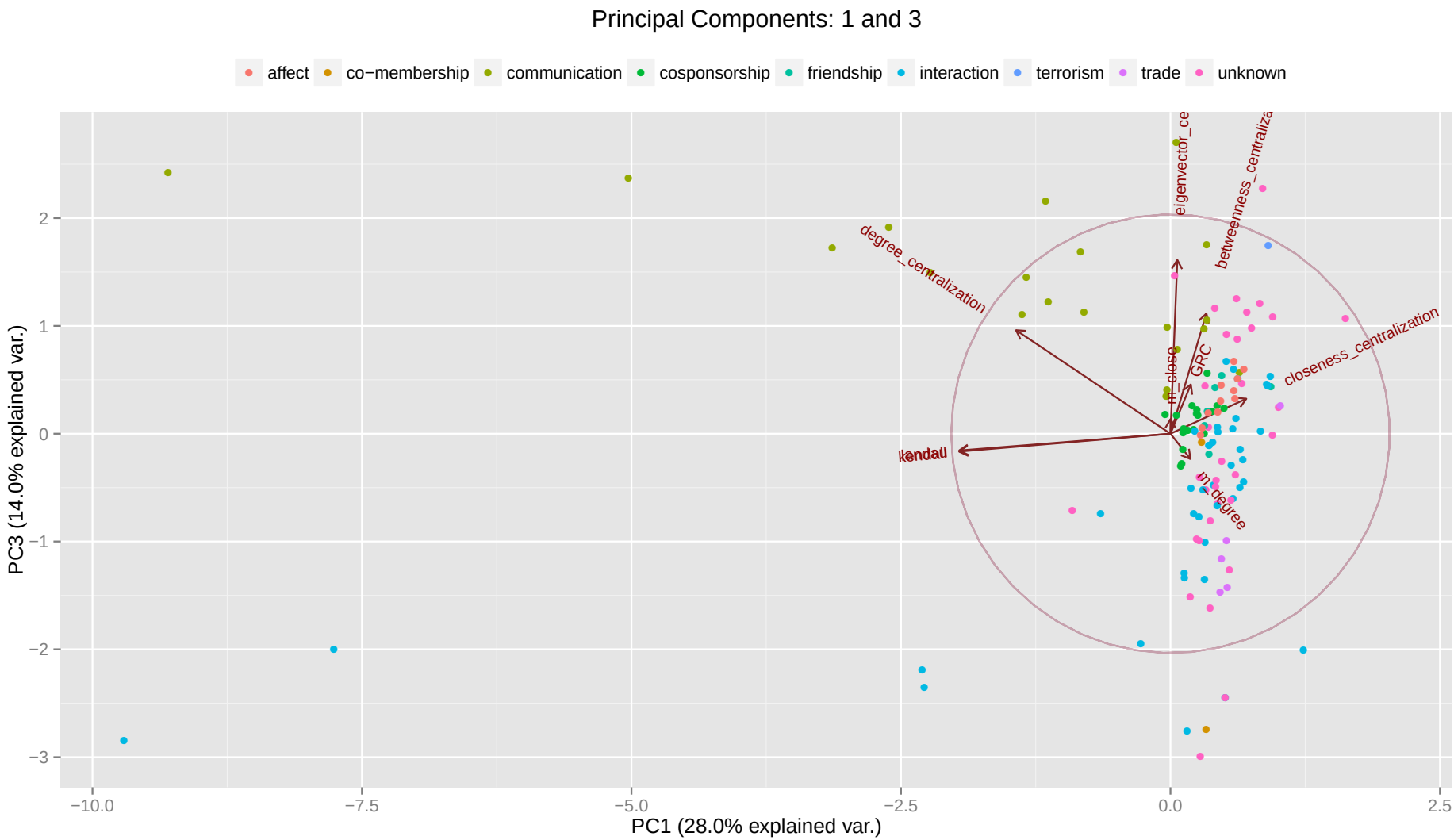
Interpretation for Components 1 and 2

- These two components explain over half of the variation
- There is a significant dropoff in the variation accounted for after PC2

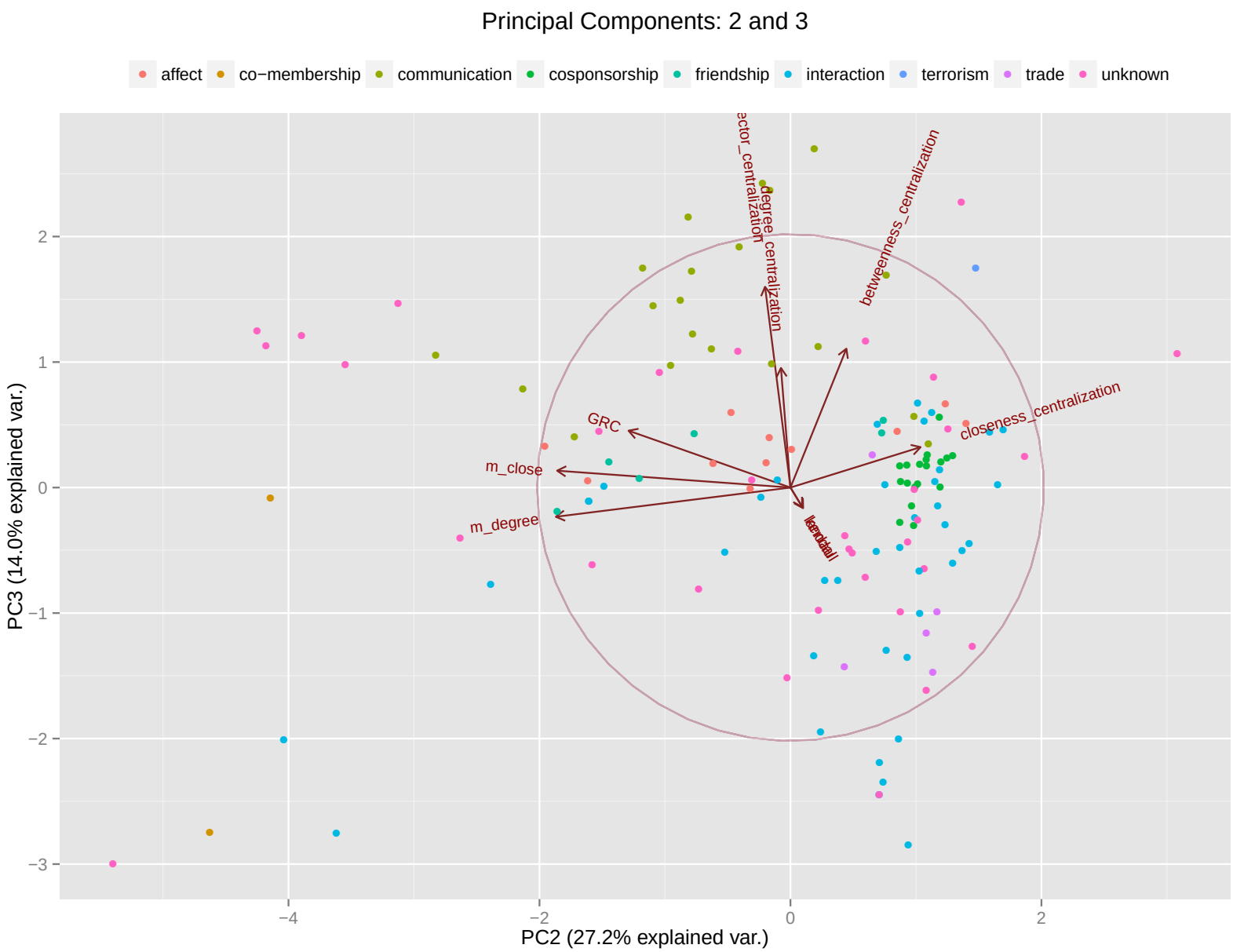
- three distinct groupings:

Control	Kendall's K, Landau's H, Degree Centralization
Reach	M Degree, M Close, GRC
Centrality	...we'll come back to this

Hierarchy in Networks – PCA



Hierarchy in Networks – PCA



Interepretation for Components 2 and 3

- Component 3 is primarily Eigenvector centralization
- the rest of component 3 is split between the various centralization measures
- back to centrality...
- the centrality measures are split evenly between the components
- we believe this is because centralization is a *related* but *distinct* concept from hierarchy

Hierarchy in Networks–Results

- Need first three principle components to explain the variation.
- Both trade and co-membership stay grouped and towards the center.
- Both Landau and Kendall stay in same direction.
- From the first PCA plot (PC1 PC2), we notice that there are three clear groupings of measures.
- we believe these correspond to two distinct models of hierarchy: dictitorial control, and decentralized reach

Hierarchy in Networks – Local Measures