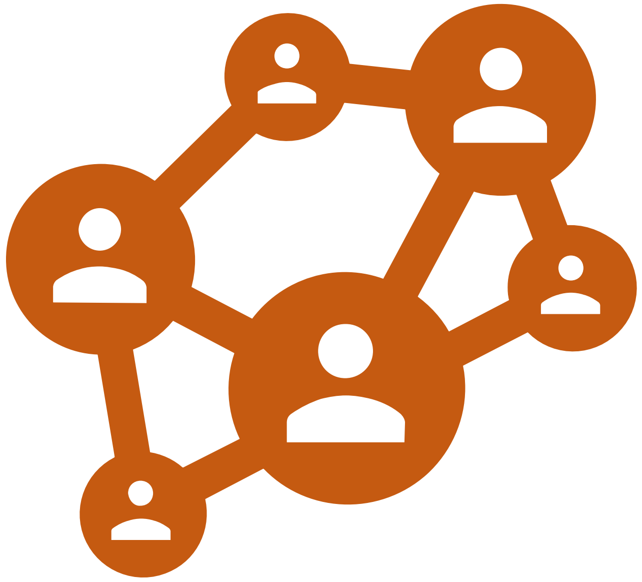
Social Network Analysis
Day 2

Advanced Winter School in Computational Social Science

Carl Nordlund
Maël Lecoursonnais
Kazuki Sakamoto



Day 2: Monday 8th Dec

- Comparing networks (QAP)
- Network null models: rewiring, configuration models, backboning
- Simulations
- Relations and nodal attributes: introduction to network dynamics
- Research design considerations

Comparing networks

- If there is a tie between two actors in one type of relation, is there likely to also be a tie between them in another relation?
  - If two countries exchange ambassadors, will they trade more?
  - Do reported friends share information with each other?
  - Is inter-personal dislike inversely related to liking each other?
  - Do we seek advice from those we trust?
- Quadratic Assignment Procedure (QAP) correlation/regression
  Knoke & Kuklinski (1982); Knoke & Wood (1981)
- For multiplex networks:
  - One set of actors/nodes
  - One or more set of relations

Knoke Bureaucracies

- Data collected by Knoke & Wood in 1978 from workers at 95 organizations in Indianapolis (Knoke & Wood 1981)
- Respondents indicated with which other organizations their own organization had any of 13 different types of relationships
- A subset of 10 organizations and two relationships (Knoke & Kuklinski (1982)
  - Money exchange (knokm.txt)
  - Information exchange (knoki.txt)

QAP

Quadratic Assignment Procedure

- Simple cell-wise correlation
- Observed correlation: -0.051
- Significant?
  - Dyadic data: observations are interdependent
- Enter QAP: permute and recalculate! Repeat N (1000+) times

knoki	COUN	COMM	EDUC	INDU	MAYR	WRO	NEWS	UWAY	WELF	WEST
COUN		1	0	0	1	0	1	0	1	0
COMM	1		1	1	1	0	1	1	1	0
EDUC	0	1		1	1	1	1	0	0	1
INDU	1	1	0		1	0	1	0	0	0
MAYR	1	1	1	1		0	1	1	1	1
WRO	0	0	1	0	0		1	0	1	0
NEWS	0	1	0	1	1	0		0	0	0
UWAY	1	1	0	1	1	0	1		1	0
WELF	0	1	0	0	1	0	1	0		0
WEST	1	1	1	0	1	0	1	0	0	

knokm	COUN	COMM	EDUC	INDU	MAYR	WRO	NEWS	UWAY	WELF	WEST
COUN		0	1	0	1	0	0	1	1	1
COMM	0		1	0	0	0	0	0	0	0
EDUC	0	0		0	0	0	0	1	0	0
INDU	0	1	1		0	0	1	1	1	0
MAYR	0	1	1	0		0	0	1	1	0
WRO	0	0	0	0	0		0	0	0	0
NEWS	0	1	0	0	0	0		1	0	0
UWAY	0	0	0	0	0	0	0		1	1
WELF	0	0	1	0	0	0	0	1		0
WEST	0	0	0	0	0	0	0	0	0	

knokm perm1	COUN	COMM	EDUC	INDU	MAYR	WRO	NEWS	UWAY	WELF	WEST
COUN		0	0	0	0	1	0	0	0	0
COMM	1		0	0	1	0	0	0	0	0
EDUC	0	0		1	1	1	0	0	1	1
INDU	0	0	0		0	0	0	0	0	0
MAYR	0	0	0	1		0	0	0	1	0
WRO	0	0	0	0	1		0	0	0	0
NEWS	0	0	0	0	0	0		0	0	0
UWAY	1	1	0	0	1	1	0		1	0
WELF	0	0	0	0	1	1	0	0		0
WEST	1	0	0	0	1	1	0	0	1	

QAP
Quadratic Assignment Procedure

- Observed cell-wise correlation: -0.051
- Average correlation from permuted data: 0.0016 [-0.466; 0.417]
- p value (probability of getting a correlation < -0.051 in permuted data)
- 0.41: i.e. quite probable, so quite insignificant

knoki	COUN	COMM	EDUC	INDU	MAYR	WRO	NEWS	UWAY	WELF	WEST
COUN		1	0	0	1	0	1	0	1	0
COMM	1		1	1	1	0	1	1	1	0
EDUC	0	1		1	1	1	1	0	0	1
INDU	1	1	0		1	0	1	0	0	0
MAYR	1	1	1	1		0	1	1	1	1
WRO	0	0	1	0	0		1	0	1	0
NEWS	0	1	0	1	1	0		0	0	0
UWAY	1	1	0	1	1	0	1		1	0
WELF	0	1	0	0	1	0	1	0		0
WEST	1	1	1	0	1	0	1	0	0	

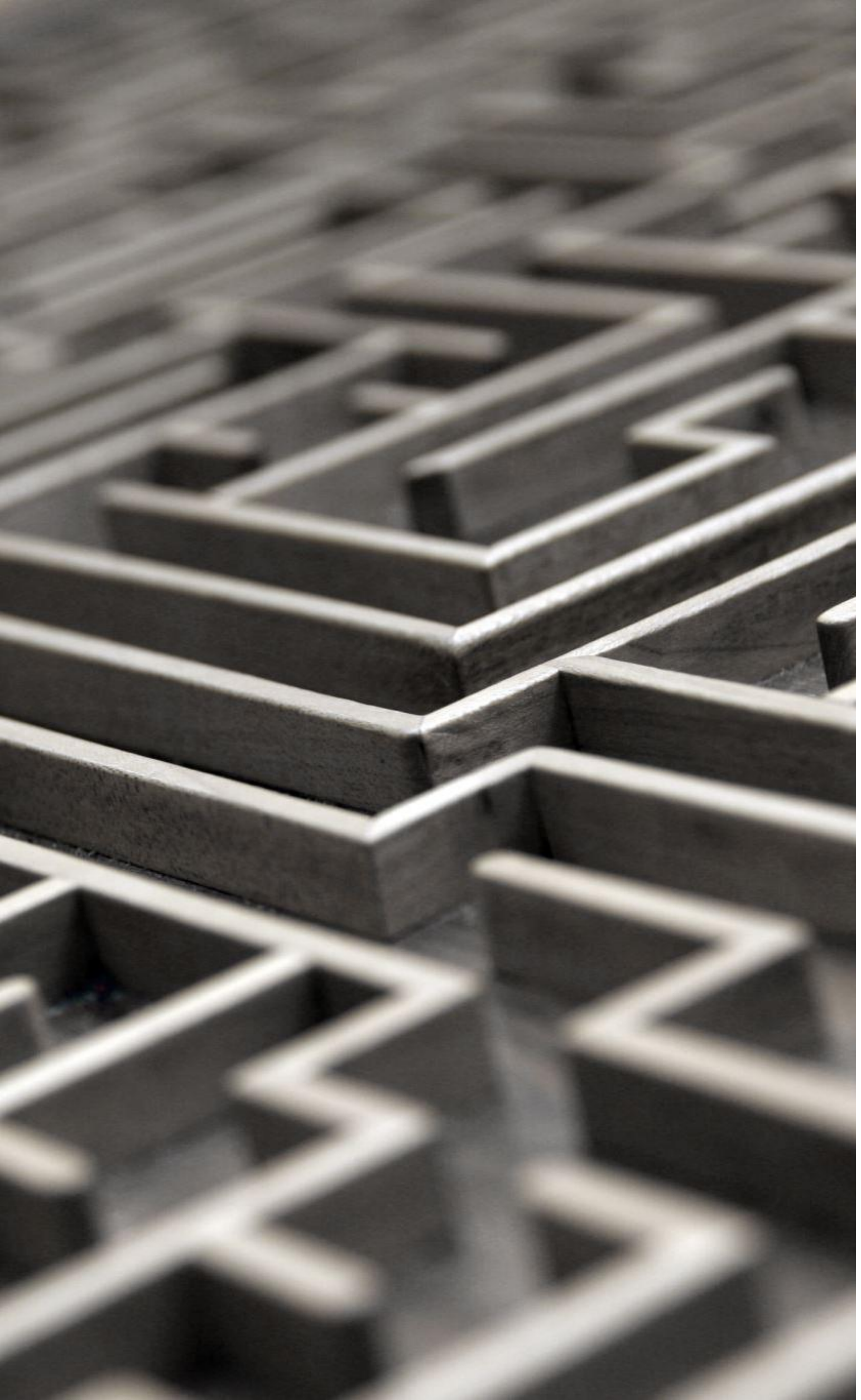
knokm	COUN	COMM	EDUC	INDU	MAYR	WRO	NEWS	UWAY	WELF	WEST
COUN		0	1	0	1	0	0	1	1	1
COMM	0		1	0	0	0	0	0	0	0
EDUC	0	0		0	0	0	0	1	0	0
INDU	0	1	1		0	0	1	1	1	0
MAYR	0	1	1	0		0	0	1	1	0
WRO	0	0	0	0	0		0	0	0	0
NEWS	0	1	0	0	0	0		1	0	0
UWAY	0	0	0	0	0	0	0		1	1
WELF	0	0	1	0	0	0	0	1		0
WEST	0	0	0	0	0	0	0	0	0	

QAP
Quadratic Assignment Procedure

- Estimating significance by random reshuffling
- Law of Large Numbers: do a lot of permutation tests
- Alternatives:
  - QAP regression
  - MR-QAP (multi-relational) regression
- Krackhardt (1988): Sampson monastery data
  - Total of 8 social relations

EIES: pre- and post-conference

- Added EIES friendship nominations at end of conference
  - eies_t2.txt
- Average friendship increased from T1 to T2
  - T1: 1.33
  - T2: 1.63
- QAP correlation: 0.809
  - Expected corr. from permutations: 0.001 [-0.221; 0.245]
  - p-value: 0.000

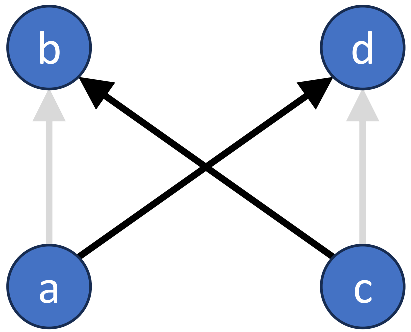


'Null' (reference/benchmark) models

- When we observe a particular network with certain features:
  - How likely is it to be as it is?
  - How to determine whether empirical observations are meaningful?
  - …or if outcomes of chance alone?
- Ways to generate reference networks (a.k.a. null models)
- Typically dyadic interdependence (as with QAP)
- Useful starting reference: Hobson et al (2021): https://onlinelibrary.wiley.com/doi/full/10.1111/brv.12775

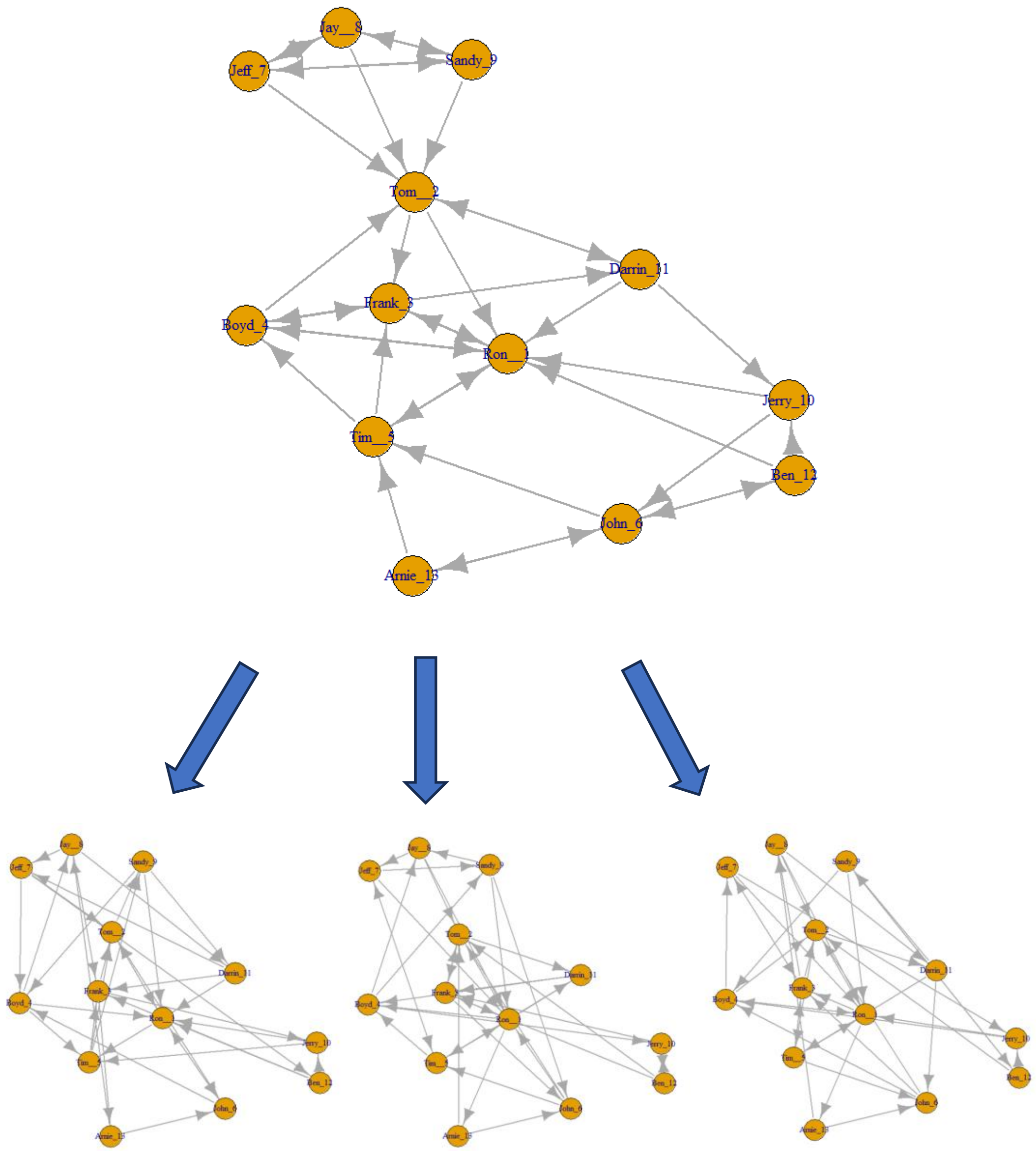
Rewiring

1. Randomly pick an edge
   a $\rightarrow$ b
2. Randomly pick another edge
   c $\rightarrow$ d
3. Make sure that:
   1. a is not connected to d
   2. c is not connected to b
4. Rewire:
   1. Remove a $\rightarrow$ b, and c $\rightarrow$ d
   2. Add a $\rightarrow$ d and c $\rightarrow$ b
5. Repeat from 1 for N (>1000) times



Rewiring

- Given one observed network
- Rewiring generates multiple random networks
- Properties of these random networks
  - Same density
  - Same degree distribution (both in- and outdegree)
- These random networks can be seen as potential counter-factuals
- Determine average macro-level properties for these to compare with observed network
- Determine assortative mixing, community structures etc.

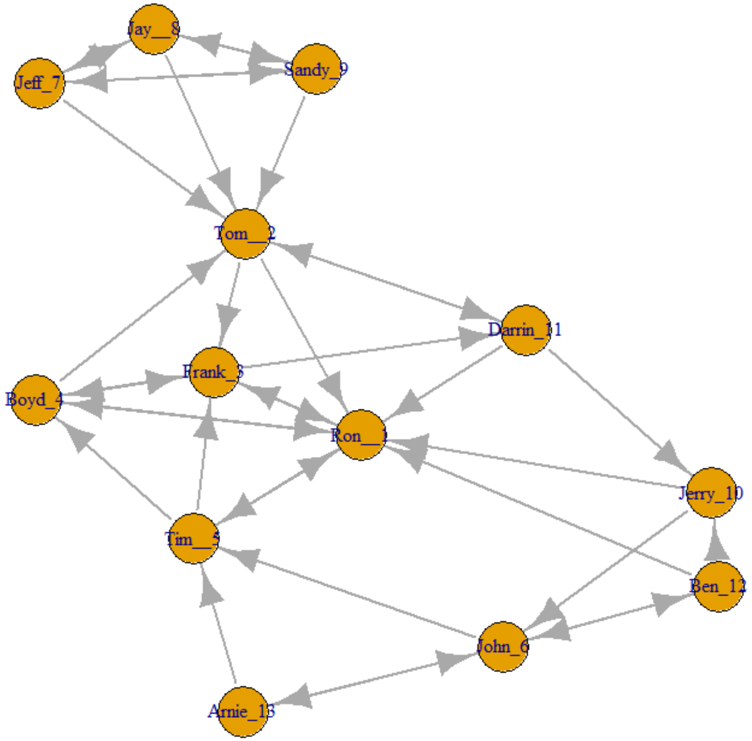


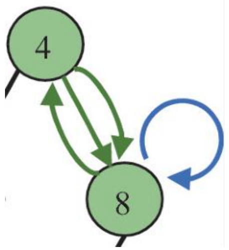
Igraph:

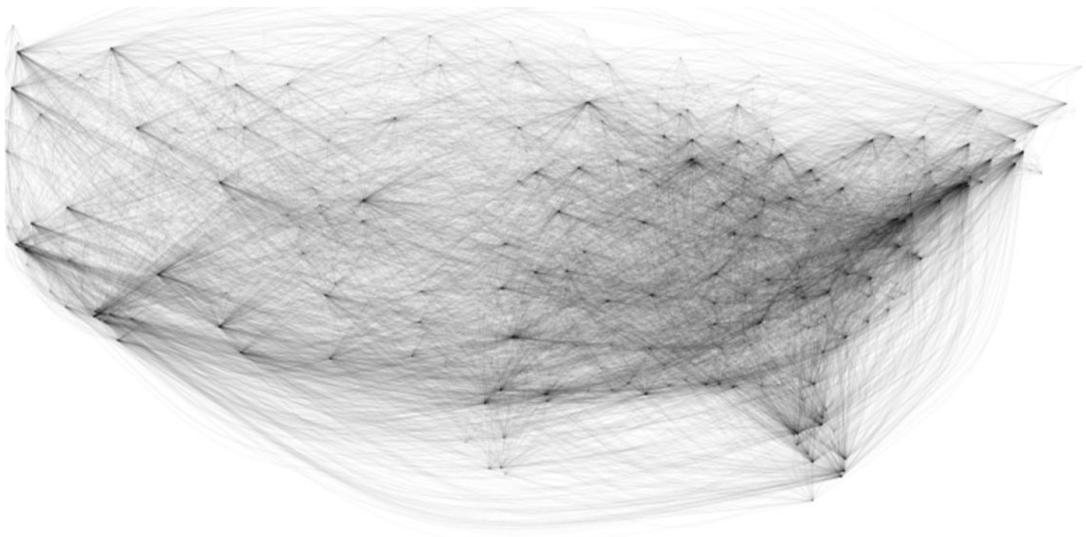
```
rewire(g,with = keeping_degseq(loops=FALSE,niter = 500))
```

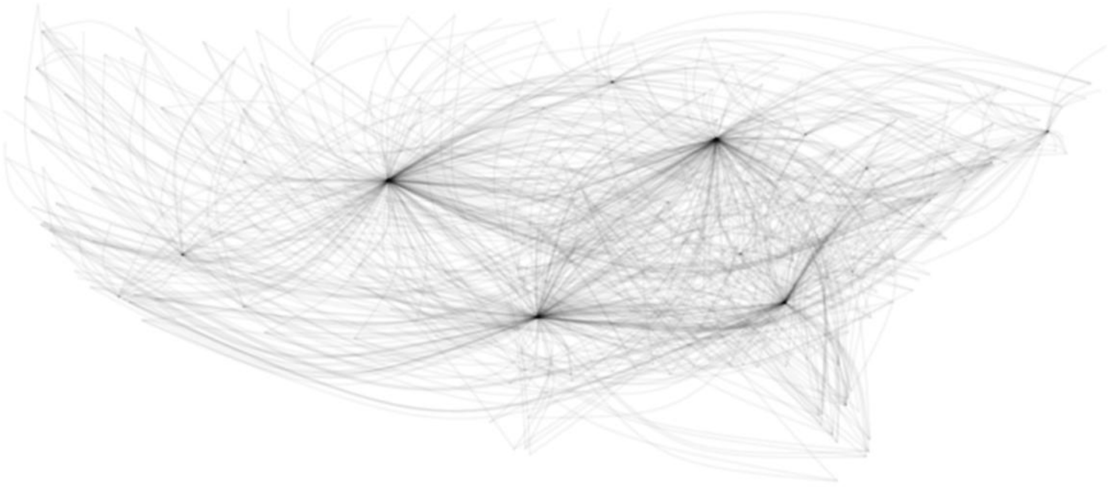
Configuration model

- Similar to rewiring
- Starts off with the degree sequence
  - I.e. assign a degree k_i to each node, typically from observed data
- Remove all edges
  - Nodes filled with stubs
  - Could be directional stubs
- Randomly pick two stubs and connect them
- Continue randomly picking stubs until none left
- Even though original network is binary and without self-ties, random network could contain both self-ties and multi-edges!









Backbone extraction

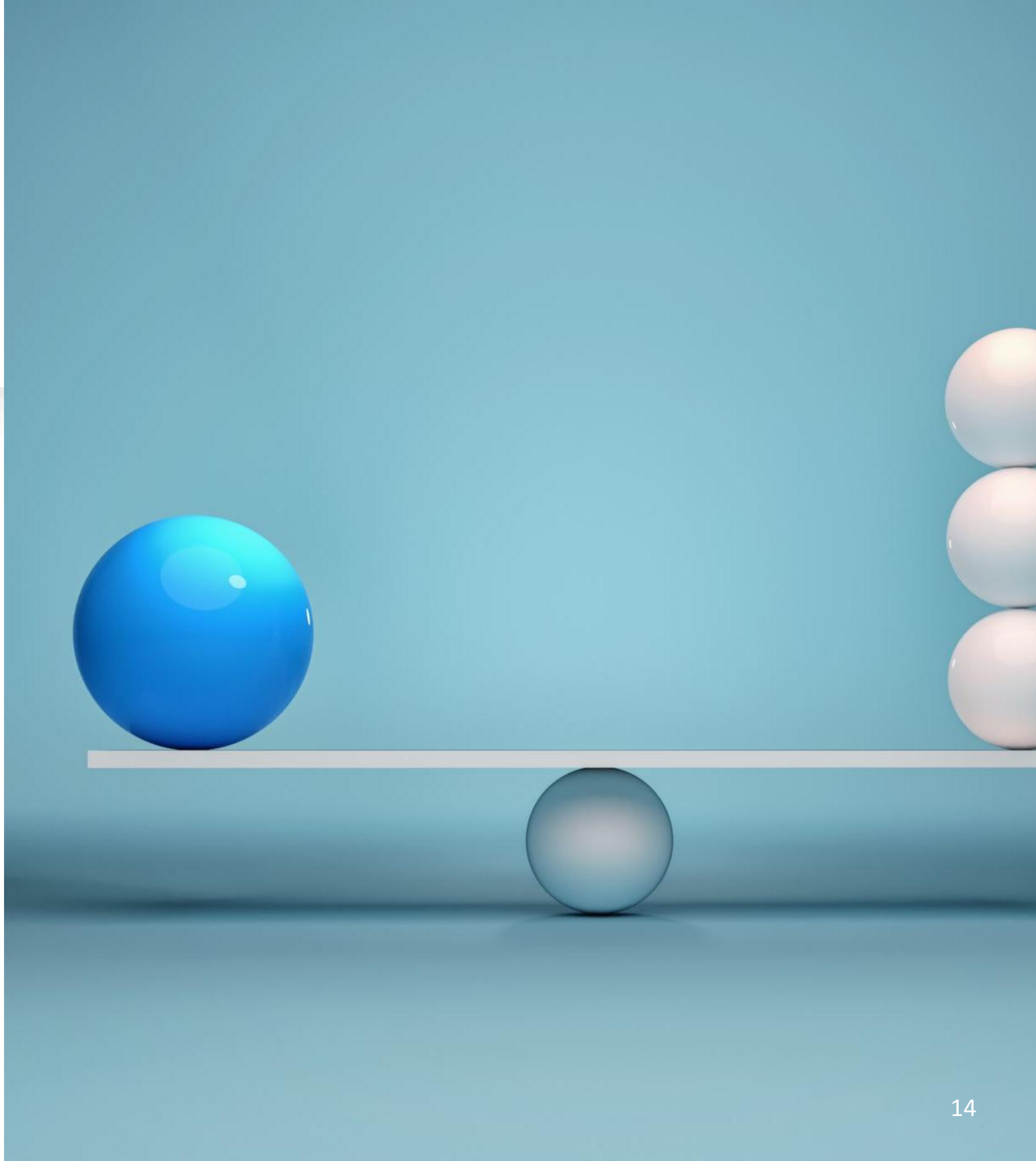
Family of methods for extracting most relevant relations in dense and/or valued networks

Also for bipartite (2-mode) networks

R package: backbone (Neal 2022)

Valued networks

- Where relations have values/weights attached to them
- Methods for valued networks
  - Blockmodeling (direct correlation-based, indirect)
  - Some centrality indices
  - Community detection (infomap)
- However, most network methods for binary networks
- Desire to convert valued networks to binary counterparts
- Rudimentary approach: dichotomization
  - But where to set the threshold?
  - What is a prominent tie?



Valued networks

	DEU	FRA	NLD	GBR	ITA	BEL	ESP	CHE	AUT	POL	SWE	CZE	NOR	IRL	HUN	DNK	PRT	SVK	FIN	ROM	GRC	SVN	LUX	BGR	LTU	EST	LVA	CYP	MLT	ISL	
DEU		103434	78104	71401	78067	62839	37030	56286	59429	37784	26758	32057	9507	4627	20990	17116	10460	10352	9040	10361	6709	4273	5220	2958	2553	1398	1282	774	391	294	761494
FRA	81804		19137	36177	42501	43135	33872	14993	4286	7504	6814	4151	2763	2442	3238	2685	5466	2337	2671	3674	3128	1432	3035	833	602	273	270	435	448	72	330178
NLD	91099	25188		40634	26426	72764	14179	7988	4281	6400	9196	4048	2888	2938	3950	5856	3866	662	3682	2181	3368	517	842	711	1028	311	441	385	22	333	336390
GBR	51108	25974	29337		13231	21871	14298	6843	2290	4676	8336	2544	4580	19488	1641	4999	2847	1134	2114	1420	1910	389	407	400	371	289	177	705	474	200	224053
ITA	57851	44962	9496	22037		11812	22127	17969	10190	9726	4366	4897	2096	1032	3739	2816	4292	2174	1917	7171	6159	4156	552	1876	764	372	371	802	1411	116	257249
BEL	44632	46993	42243	26347	17683		7933	4927	2427	4104	5709	2276	1418	1436	1927	2736	2150	679	1620	1313	2167	385	5069	482	756	187	193	177	91	53	228113
ESP	29486	37185	9330	15494	22153	8151		4724	2348	3490	1818	2253	1015	882	1097	1183	23557	647	897	1303	1924	705	239	479	281	123	132	238	144	45	171323
CHE	43552	14600	2809	9185	13540	3855	3576		8138	1499	1247	1353	802	1123	663	852	480	392	774	584	1043	515	152	235	61	143	162	104	143	54	111636
AUT	45234	5876	2334	4056	11197	2442	2460	7625		2914	1603	4231	614	219	5414	775	383	1630	625	2539	706	2091	127	882	172	94	129	39	24	15	106450
POL	37648	9124	6130	9388	9570	3765	4034	1239	2512		4342	8041	1941	425	4590	2463	466	2632	1200	2318	482	574	177	533	2065	718	878	28	26	48	117357
SWE	17526	7474	7002	10082	4590	7325	3064	1366	1653	3263		1168	10807	452	794	11009	768	344	6907	352	472	178	107	133	763	1121	396	60	20	204	99400
CZE	39246	6640	6024	6141	5938	3110	2929	2049	5547	6442	1789		767	255	2832	951	470	6643	802	1475	311	633	139	482	332	141	163	30	12	22	102315
NOR	22663	6491	11177	29730	2073	4623	2412	276	769	2671	12971	1037		1712	29	3306	700	62	1573	117	263	39	35	11	79	238	117	9	2	354	105539
IRL	18591	7745	5295	19745	4153	19810	4341	5629	805	1092	1943	1052	798		457	974	725	184	509	323	551	95	84	71	42	57	27	234	39	32	95403
HUN	22126	4233	2855	5016	4763	1459	2388	891	4150	3057	991	2734	348	268		588	353	2758	335	5383	301	757	115	791	187	92	148	31	19	13	67150
DNK	14700	3524	3612	6312	2808	1484	2222	937	601	2115	12309	772	4801	1032	649		416	272	1635	270	644	78	53	127	398	181	255	36	111	276	62630
PRT	5504	5579	1909	2667	1837	1439	11319	484	529	443	510	401	257	131	148	367		118	402	251	146	46	75	74	18	16	12	45	9	8	34744
SVK	12341	3841	1652	2492	3383	955	2105	491	3485	3599	1122	6505	393	69	3624	339	135		278	959	169	371	77	276	139	40	80	19	4	6	48949
FIN	7979	2923	4211	3339	1921	2206	1410	812	588	1765	7814	450	1973	252	434	1370	163	162		233	345	90	25	126	414	1521	547	30	27	207	43337
ROM	8860	3973	1249	1907	6185	720	1382	295	1248	1225	307	708	477	141	2281	125	159	462	134		607	252	17	1769	42	24	8	91	18	5	34671
GRC	2592	811	523	1040	2490	336	713	214	182	304	197	175	43	39	107	184	142	122	182	827		105	8	1511	17	9	10	1616	77	3	14579
SVN	5035	1829	382	547	2868	208	291	247	1664	683	199	531	97	17	856	235	46	247	97	413	185		22	200	75	25	35	9	4	3	17050
LUX	3945	1983	859	1441	1399	2761	594	248	246	242	355	215	84	36	105	139	66	59	78	63	208	72		17	13	7	16	6	1	1	15259
BGR	2302	857	259	353	2144	1002	474	102	427	333	90	176	46	14	202	69	35	99	46	1900	1375	159	4		51	7	19	42	4	2	12593
LTU	1955	790	1102	851	372	323	254	51	83	1022	787	119	529	56	67	527	35	24	300	46	33	9	1	38		790	1900	2	1	18	12085
EST	583	191	325	237	114	149	52	85	48	166	1682	59	423	4	17	238	12	18	1664	14	6	6	2	5	668		797	3	1	33	7602
LVA	802	173	228	588	125	84	107	31	31	220	578	40	257	33	20	310	5	21	284	31	9	3	1	15	1461	705		2	1	16	6181
CYP	348	46	146	160	222	86	16	5	47	34	27	39	6	4	77	31	2	9	7	83	719	1	5	37	4	2	22		8	0	2193
MLT	404	481	116	257	332	104	139	11	13	45	14	15	9	15	21	13	30	4	21	29	38	3	2	34	1	1	1	134		7	2294
ISL	940	150	834	546	53	89	169	64	21	93	34	58	208	13	9	101	18	14	14	4	10	12	2	1	13	3	4	0	1		3478
	670856	373070	248680	328170	282138	278907	175890	136882	118038	106911	113908	82105	49947	39155	59978	62357	58247	34261	39808	45637	33988	17946	16594	15107	13370	8888	8592	6086	3739	2440	

Iceland's largest import (333) is smaller than Germany's smallest import (348)

- Similar relational capacities?
- Likert-scales
  - Friendship nominations (EIES)
  - Notesharing (Hlebec)
- Allocation-style
  - Minutes spent during lunch break
  - Rank ordering
- Non-restrained
  - International trade flows

Valued networks
Unequal relational capacities

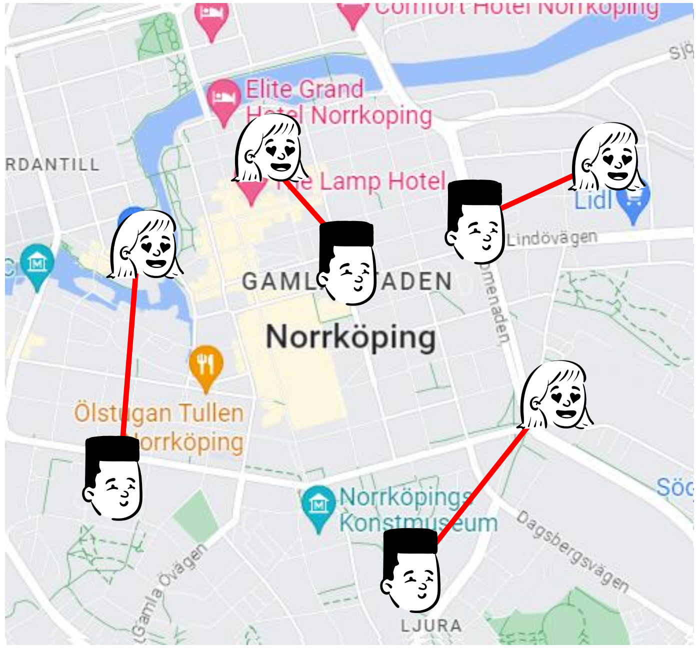
	DEU	FRA	NLD	GBR	ITA	BEL	ESP	CHE	AUT	POL	SWE	CZE	NOR	IRL	HUN	DNK	PRT
DEU		103434	78104	71401	78067	62839	37030	56286	59429	37784	26758	32057	9507	4627	20990	17116	10460
FRA	81804		19137	36177	42501	43135	33872	14993	4286	7504	6814	4151	2763	2442	3238	2685	5466
NLD	91099	25188		40634	26426	72764	14179	7988	4281	6400	9196	4048	2888	2938	3950	5856	3866
GBR	51108	25974	29337		13231	21871	14298	6843	2290	4676	8336	2544	4580	19488	1641	4999	2847
ITA	57851	44962	9496	22037		11812	22127	17969	10190	9726	4366	4897	2096	1032	3739	2816	4292
BEL	44632	46993	42243	26347	17683		7933	4927	2427	4104	5709	2276	1418	1436	1927	2736	2150
ESP	29486	37185	9330	15494	22153	8151		4724	2348	3490	1818	2253	1015	882	1097	1183	23557
CHE	43552	14600	2809	9185	13540	3855	3576		8138	1499	1247	1353	802	1123	663	852	480
AUT	45234	5876	2334	4056	11197	2442	2460	7625		2914	1603	4231	614	219	5414	775	383
POL	37648	9124	6130	9388	9570	3765	4034	1239	2512		4342	8041	1941	425	4590	2463	466
SWE	17526	7474	7002	10082	4590	7325	3064	1366	1653	3263		1168	10807	452	794	11009	768
CZE	39246	6640	6024	6141	5938	3110	2929	2049	5547	6442	1789		767	255	2832	951	470
NOR	22663	6491	11177	29730	2073	4623	2412	276	769	2671	12971	1037		1712	29	3306	700
IRL	18591	7745	5295	19745	4153	19810	4341	5629	805	1092	1943	1052	798		457	974	725
HUN	22126	4233	2855	5016	4763	1459	2388	891	4150	3057	991	2734	348	268		588	353
DNK	14700	3524	3612	6312	2808	1484	2222	937	601	2115	12309	772	4801	1032	649		416
PRT	5504	5579	1909	2667	1837	1439	11319	484	529	443	510	401	257	131	148	367	
SVK	12341	3841	1652	2492	3383	955	2105	491	3485	3599	1122	6505	393	69	3624	339	135
FIN	7979	2923	4211	3339	1921	2206	1410	812	588	1765	7814	450	1973	252	434	1370	163
ROM	8860	3973	1249	1907	6185	720	1382	295	1248	1225	307	708	477	141	2281	125	159
GRC	2592	811	523	1040	2490	336	713	214	182	304	197	175	43	39	107	184	142
SVN	5035	1829	382	547	2868	208	291	247	1664	683	199	531	97	17	856	235	46
LUX	3945	1983	859	1441	1399	2761	594	248	246	242	355	215	84	36	105	139	66
BGR	2302	857	259	353	2144	1002	474	102	427	333	90	176	46	14	202	69	35
LTU	1955	790	1102	851	372	323	254	51	83	1022	787	119	529	56	67	527	35
EST	583	191	325	237	114	149	52	85	48	166	1682	59	423	4	17	238	12
LVA	802	173	228	588	125	84	107	31	31	220	578	40	257	33	20	310	5
CYP	348	46	146	160	222	86	16	5	47	34	27	39	6	4	77	31	2
MLT	404	481	116	257	332	104	139	11	13	45	14	15	9	15	21	13	30
ISL	940	150	834	546	53	89	169	64	21	93	34	58	208	13	9	101	18
	670856	373070	248680	328170	282138	278907	175890	136882	118038	106911	113908	82105	49947	39155	59978	62357	58247

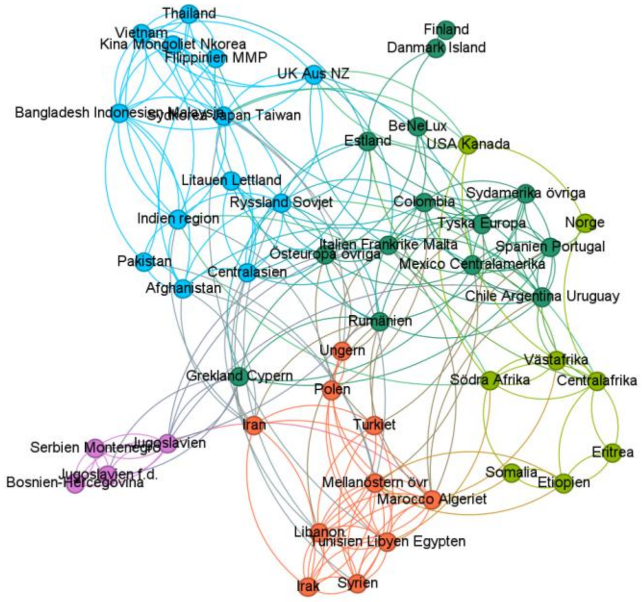
	SVK	FIN	ROM	GRC	SVN	LUX	BGR	LTU	EST	LVA	CYP	MLT	ISL	
DEU	10352	9040	10361	6709	4273	5220	2958	2553	1398	1282	774	391	294	761494
FRA	2337	2671	3674	3128	1432	3035	833	602	273	270	435	448	72	330178
NLD	662	3682	2181	3368	517	842	711	1028	311	441	385	228	333	336390
GBR	1134	2114	1420	1910	389	407	400	371	289	177	705	474	200	224053
ITA	2174	1917	7171	6159	4156	552	1876	764	372	371	802	1411	116	257249
BEL	679	1620	1313	2167	385	5069	482	756	187	193	177	91	53	228113
ESP	647	897	1303	1924	705	239	479	281	123	132	238	144	45	171323
CHE	392	774	584	1043	515	152	235	61	143	162	104	143	54	111636
AUT	1630	625	2539	706	2091	127	882	172	94	129	39	24	15	106450
POL	2632	1200	2318	482	574	177	533	2065	718	878	28	26	48	117357
SWE	344	6907	352	472	178	107	133	763	1121	396	60	20	204	99400
CZE	6643	802	1475	311	633	139	482	332	141	163	30	12	22	102315
NOR	62	1573	117	263	39	35	11	79	238	117	9	2	354	105539
IRL	184	509	323	551	95	84	71	42	57	27	234	39	32	95403
HUN	2758	335	5383	301	757	115	791	187	92	148	31	19	13	67150
DNK	272	1635	270	644	78	53	127	398	181	255	36	111	276	62630
PRT	118	402	251	146	46	75	74	18	16	12	45	9	8	34744
SVK		278	959	169	371	77	276	139	40	80	19	4	6	48949
FIN	162		233	345	90	25	126	414	1521	547	30	27	207	43337
ROM	462	134		607	252	17	1769	42	24	8	91	18	5	34671
GRC	122	182	827		105	8	1511	17	9	10	1616	77	3	14579
SVN	247	97	413	185		22	200	75	25	35	9	4	3	17050
LUX	59	78	63	208	72		17	13	7	16	6	1	1	15259
BGR	99	46	1900	1375	159	4		51	7	19	42	4	2	12593
LTU	24	300	46	33	9	1	38		790	1900	2	1	18	12085
EST	18	1664	14	6	6	2	5	668		797	3	1	33	7602
LVA	21	284	31	9	3	1	15	1461	705		2	1	16	6181
CYP	9	7	83	719	1	5	37	4	2	22		8	0	2193
MLT	4	21	29	38	3	2	34	1	1	1	134		7	2294
ISL	14	14	4	10	12	2	1	13	3	4	0	1		3478
	34261	39808	45637	33988	17946	16594	15107	13370	8888	8592	6086	3739	2440	

- Transaction flow model (Savage & Deutsch 1960)
    - Relative acceptance index
    - Contingency table approaches
- Iterative row-column normalization (e.g. Schwartz 1977)
    - Sums of rows and columns converge to unity
    - Dichotomize these
- Deviational transformation (Nordlund & Žiberna 2020)
- Determine baseline model for valued networks
- Positive deviations from such baselines indicate prominent (binary) ties

Simulation-based null models

- When tie formation mechanisms and behaviors are reasonably understood
- Using custom-made simulations do generate counterfactual networks
- Determinants for tie formation and rewiring
  - Contextual knowledge/insights
  - Theoretical considerations
  - Attributes of observed actors





Example:
Alternative family formation events

(Lindmarker & Nordlund, ongoing)

Given existing set of geolocated heterosexual couples (prior to moving in together)

For each individual, compile a set of potential partners
- Centered around the X,Y coordinate of actual partner
- Potential partner ages: +/- 2 years of actual partner (reciprocated)
- Overlapping activity period (5 years)

Generate new families

Collapse into intra- and inter-group networks

Generate 100+ simulated networks

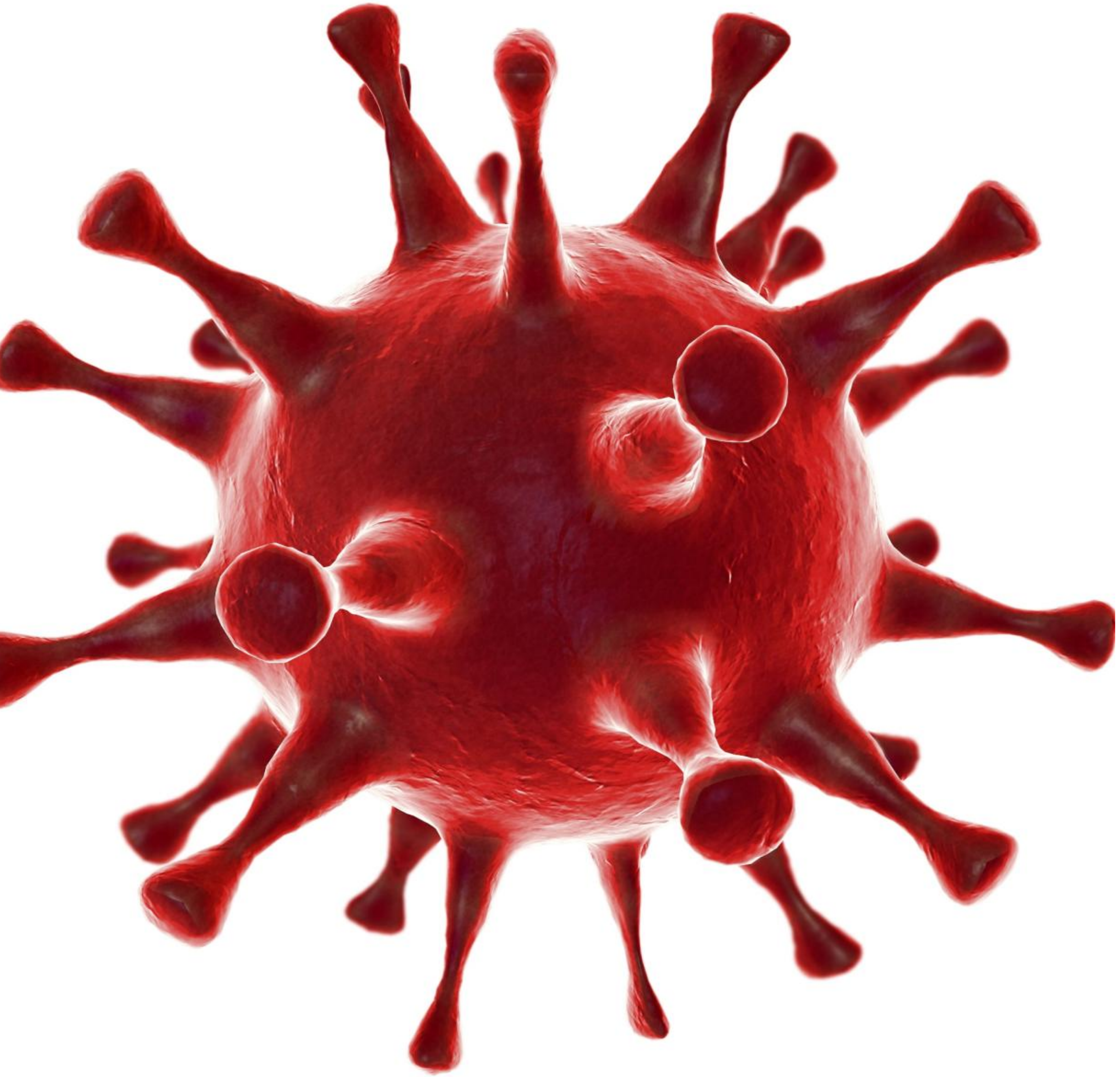
Compare with original observed network

Relations and nodal attributes: Introduction to network dynamics
(stolen from David Garcia)

- Assortativity
  - Similar nodes more likely to be connected than dissimilar nodes
  - Assortativity coefficient
- But what generates assortativity?
  - Homophily
  - Contagion
  - Confounding

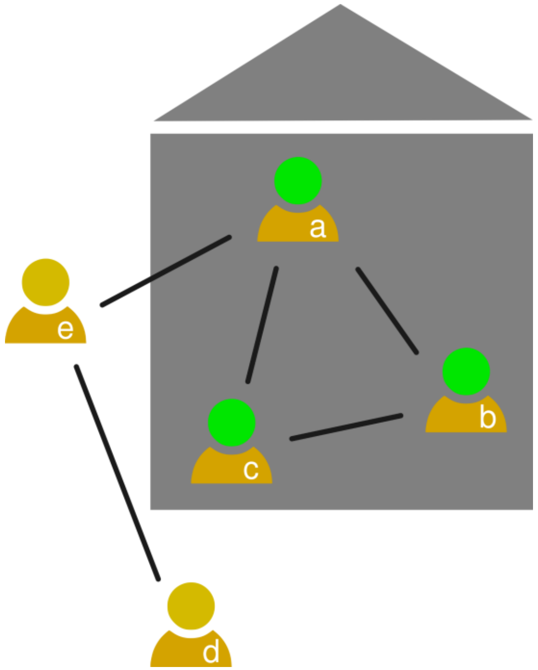
Homophily

- Tie formation
  - Individuals more likely to form social ties to individuals that are similar
  - “Similarity breeds connection”
- Attributes (e.g. gender, opinions etc) begets ties
- Selection
- Birds of a feather flock together (McPherson et al 2001)



Contagion

- Change of attributes
  - Individuals are influenced by their social surroundings
  - Adopting behaviors of their social contacts
- Relations beget attributes
- Influence
- Contagion (e.g. Centola 2007; 2018)
  - Simple (epidemic-style)
  - Complex (needs more than one alter influence)



Confounding

- External effects can produce both tie formation and attributional change
- No causal relationship between relations and attributes
- A third cause!

Disentangling choice vs. influence

- Typically requires more than a simple snapshot
- Longitudinal data preferable

Social Network Analysis and CSS

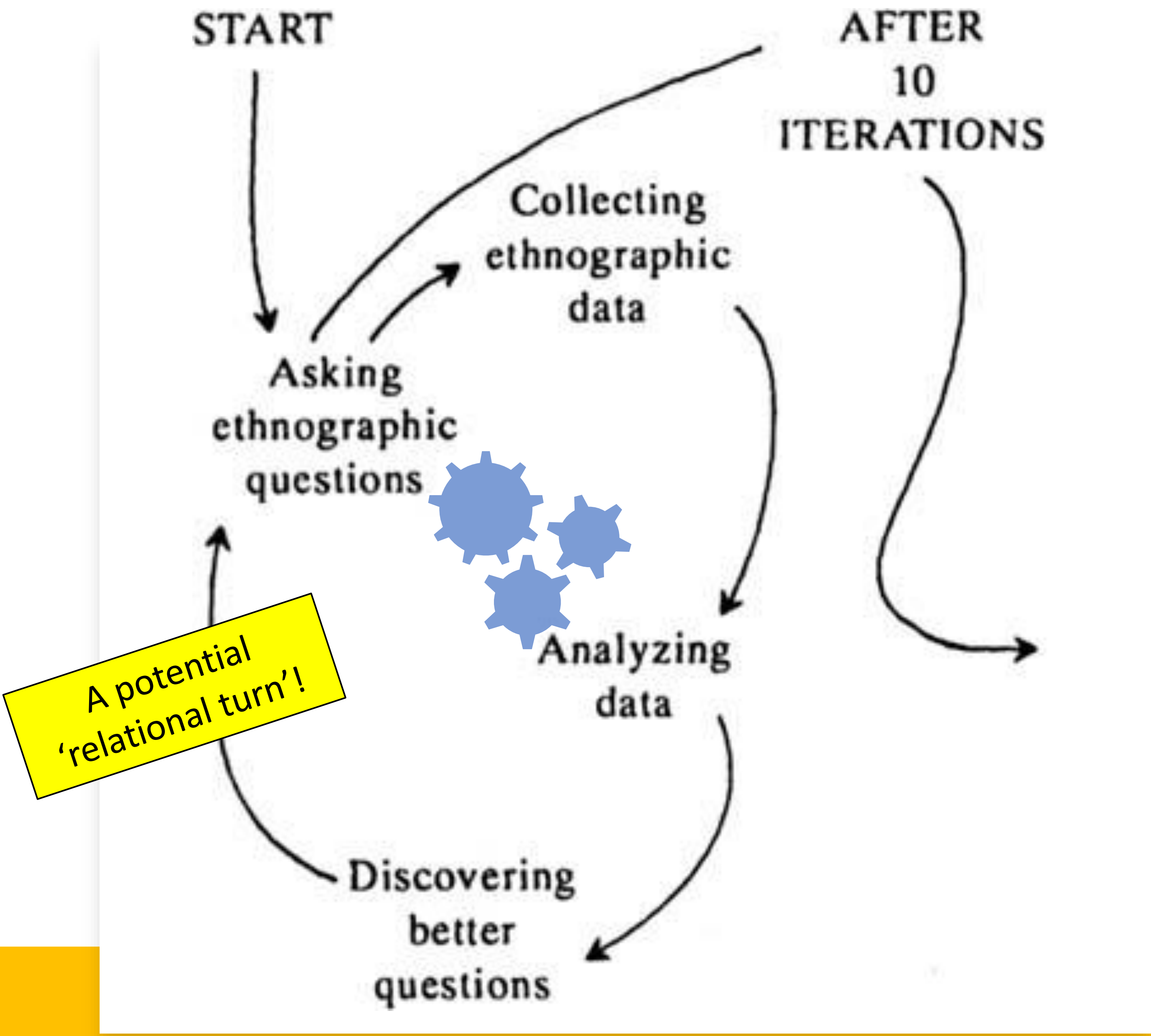
- In CSS, we tend to start off with the RQ cog
- Building on previous research, we formulate RQs of social-scientific and/or societal relevance
    - Is GDP growth linked to positionality in trade networks?
    - Are central NGO's more efficient than less central ones?
    - Are inter-ethnic families happier than endogamic ones?
- To address these: pair up methods with data
- Circling back to RQ at the end

Social Network Analysis and CSS

- Postulating that data science is more exploratory
- Given (typically large and complex) datasets
  - More exploratory, open-minded approach
  - Apply methods to prune out insights data
- Although general topic area given by the data
  - Perhaps more open to *finding* signals
  - Seed to an interesting research question
- Formulating question (or topic)
- Connecting with previous research

Social Network Analysis and CSS

- Ethnographic cycle (Gladwin 1989)
    1. Ask ethnographic questions
    2. Collect data
    3. Analyze data
    4. Discover better question (repeat from 1)
- In context of CSS (and SNA)
    1. Research question
    2. Database, scraping, wrangling
    3. Methods and analysis
    4. Discover better questions (repeat from 1)



THANK YOU FOR YOUR ATTENTION TO THIS MATTER!

carlnordlund.net