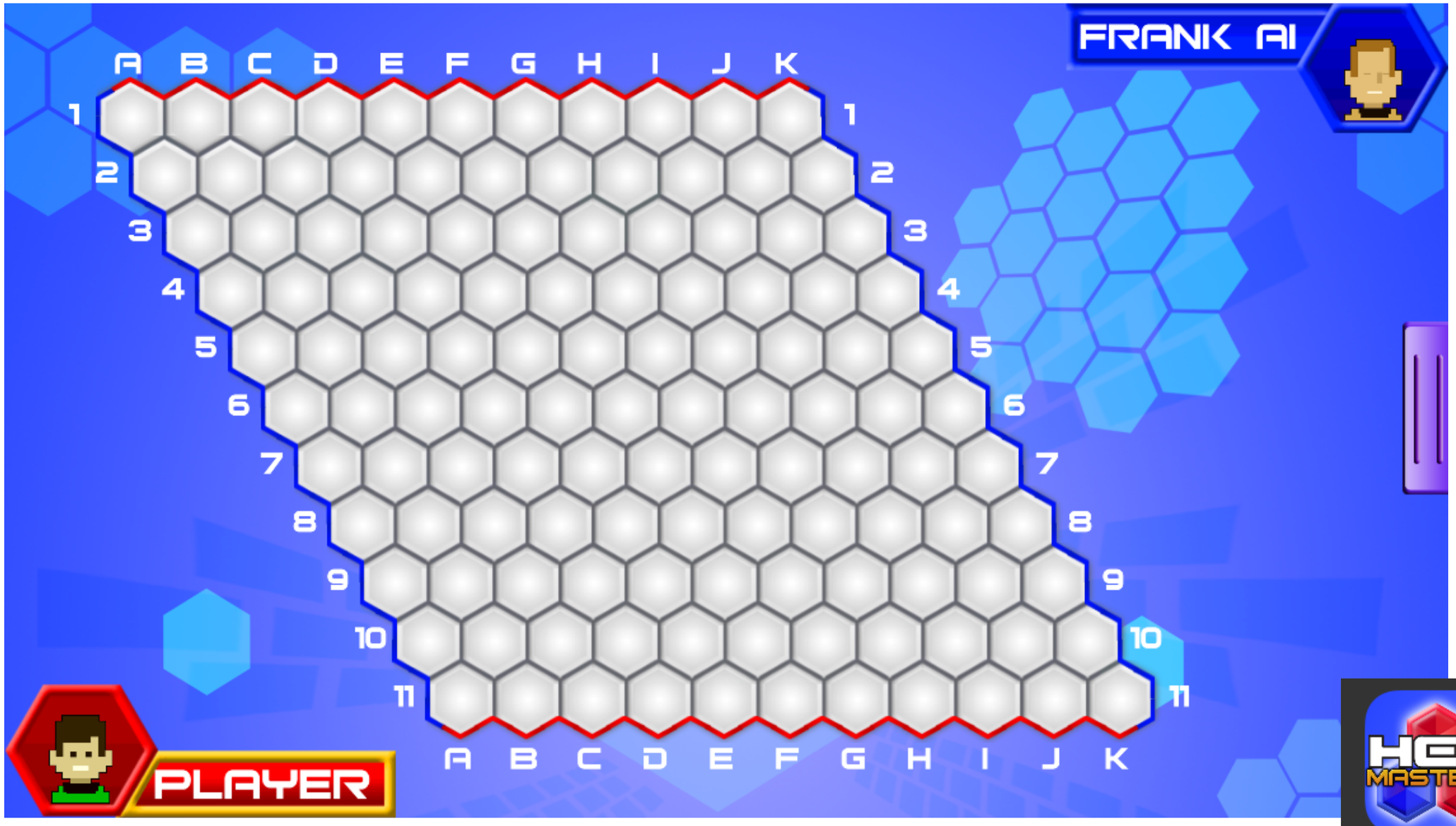
Lecture 2

Review


FRANK AI
A B C D E F G H I J K
1 2 3 4 5 6 7 8 9 10 11
PLAYER
Hex Masters

What this course is about...

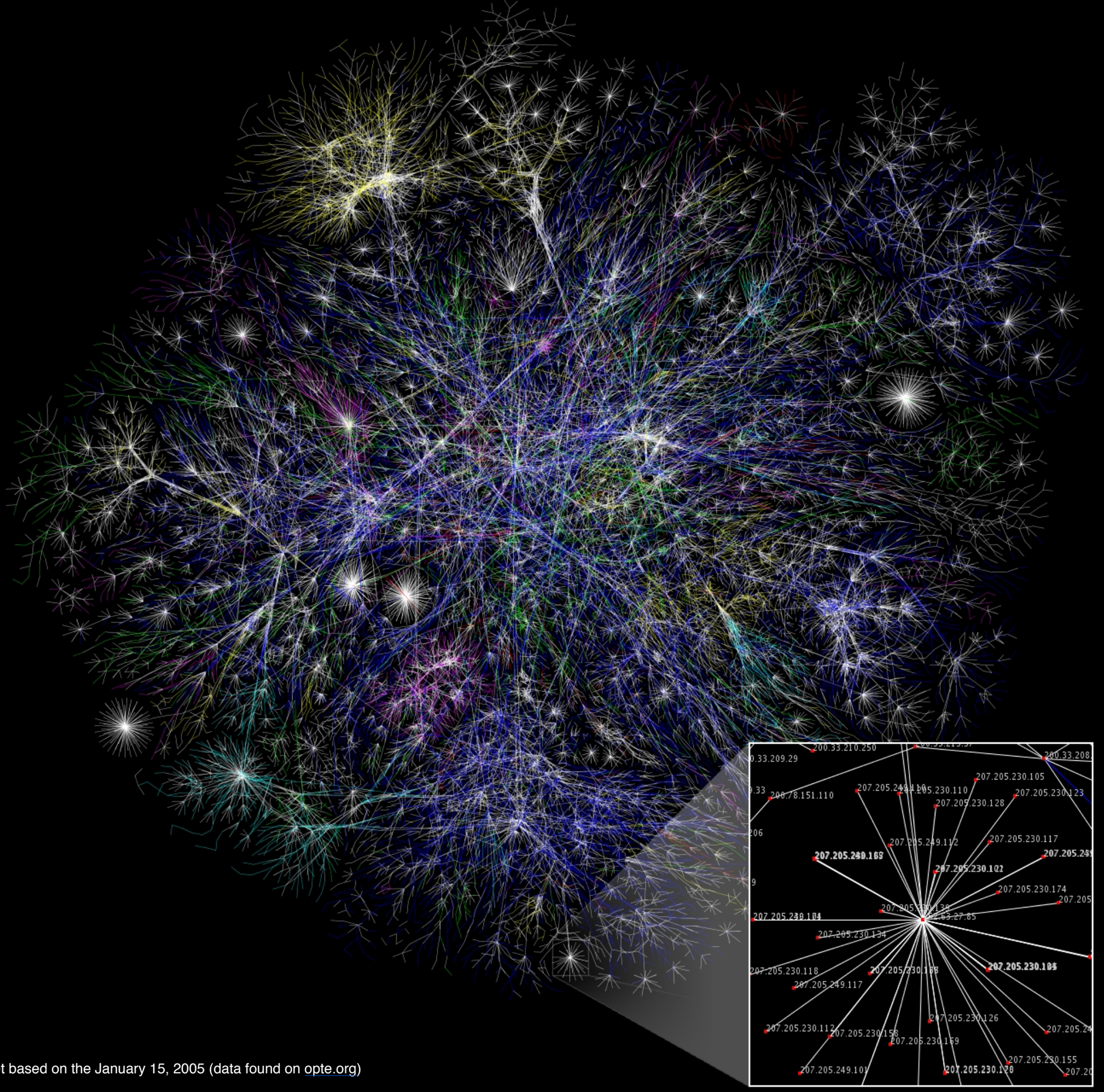
- Tools for research
- Theoretical foundations, mathematical tools, modeling, equilibrium notions

Focus

- Graph theory: model network structure
- Game theory: model behavior where decisions affect each other's behavior

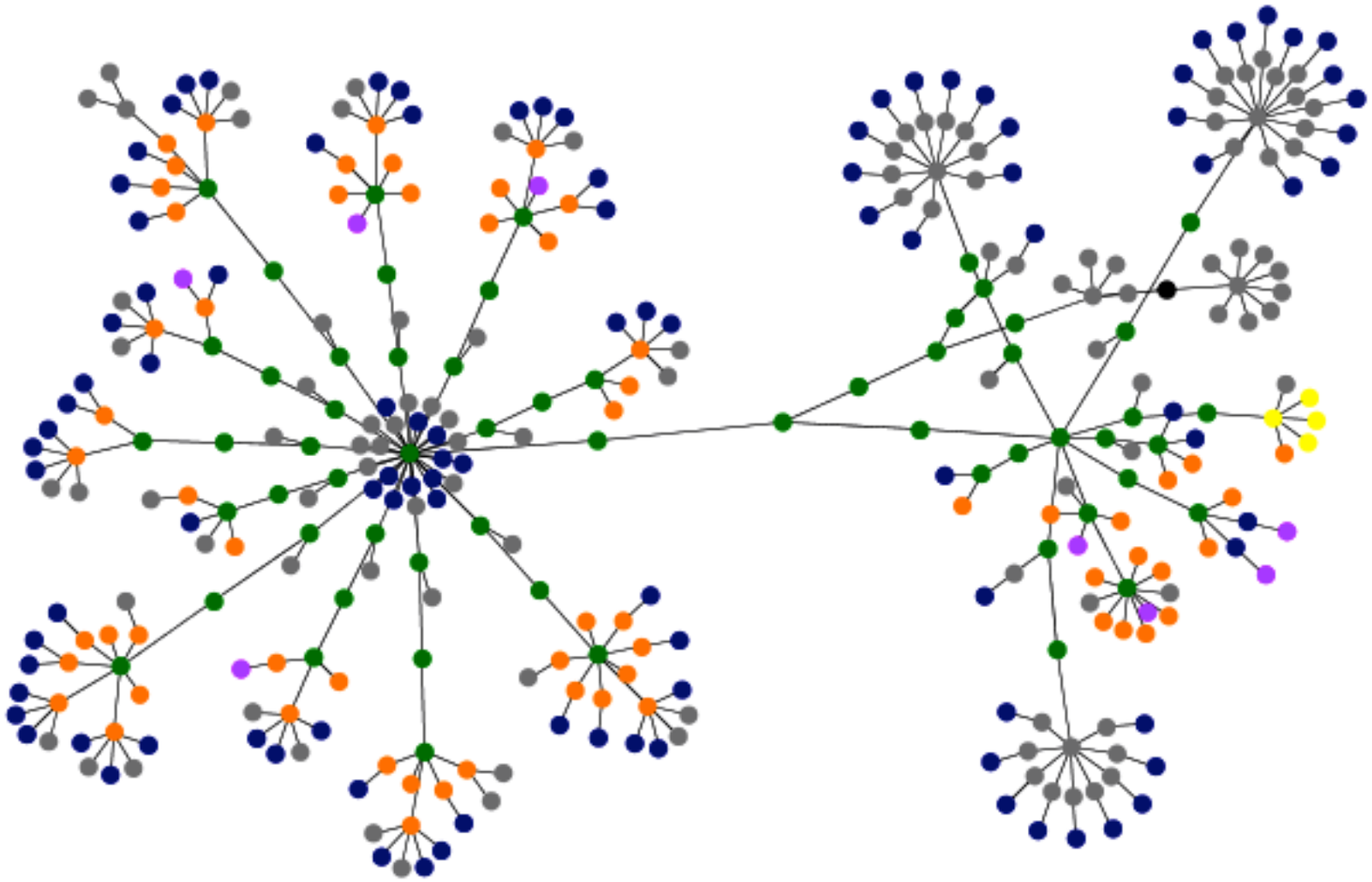
Related courses

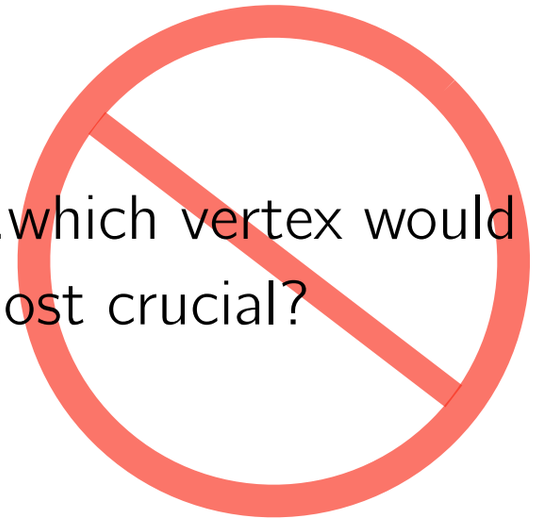
- Behavioral Game Theory
- Information Networks



Partial map of the Internet based on the January 15, 2005 (data found on opte.org)

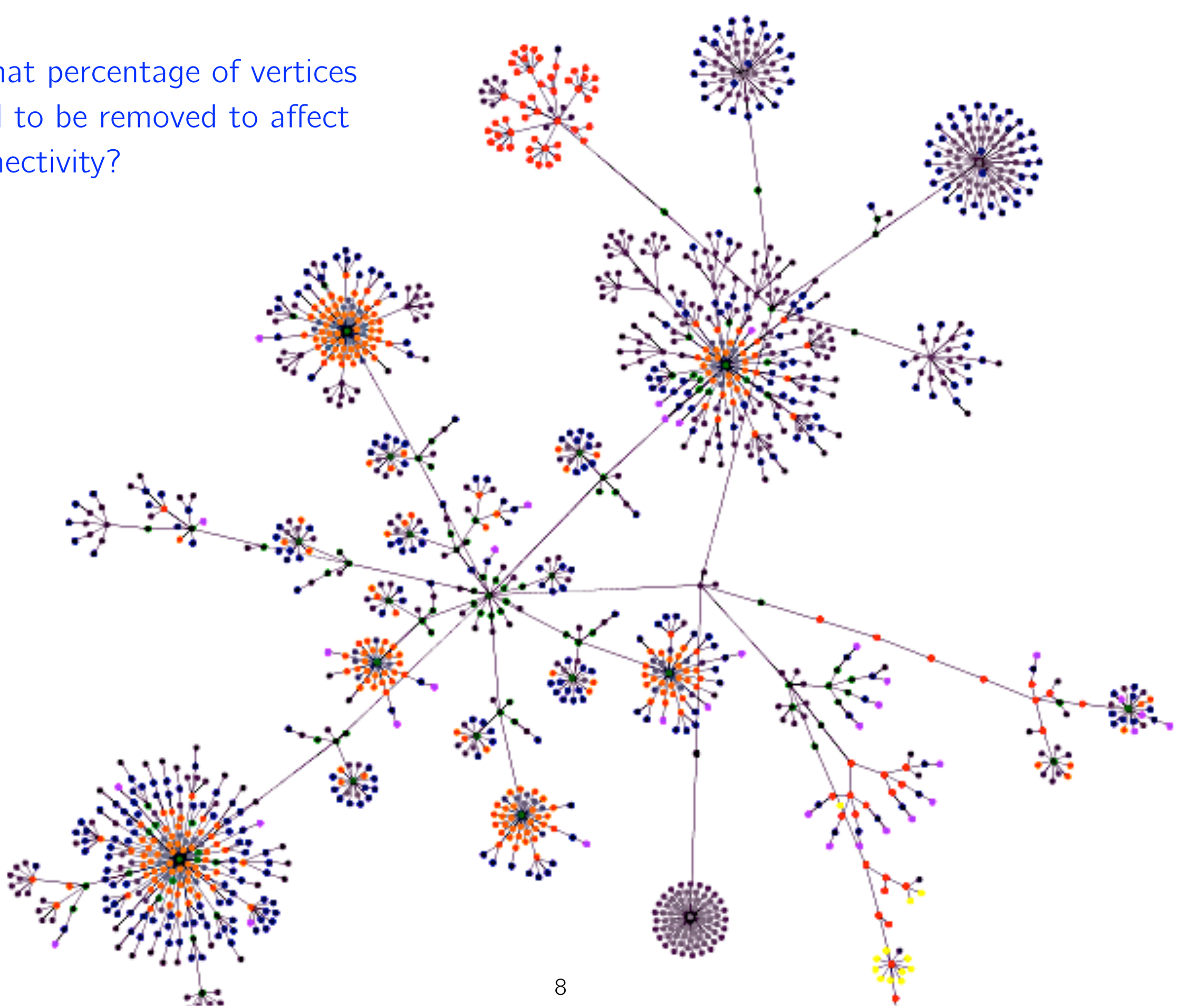
As networks grow...



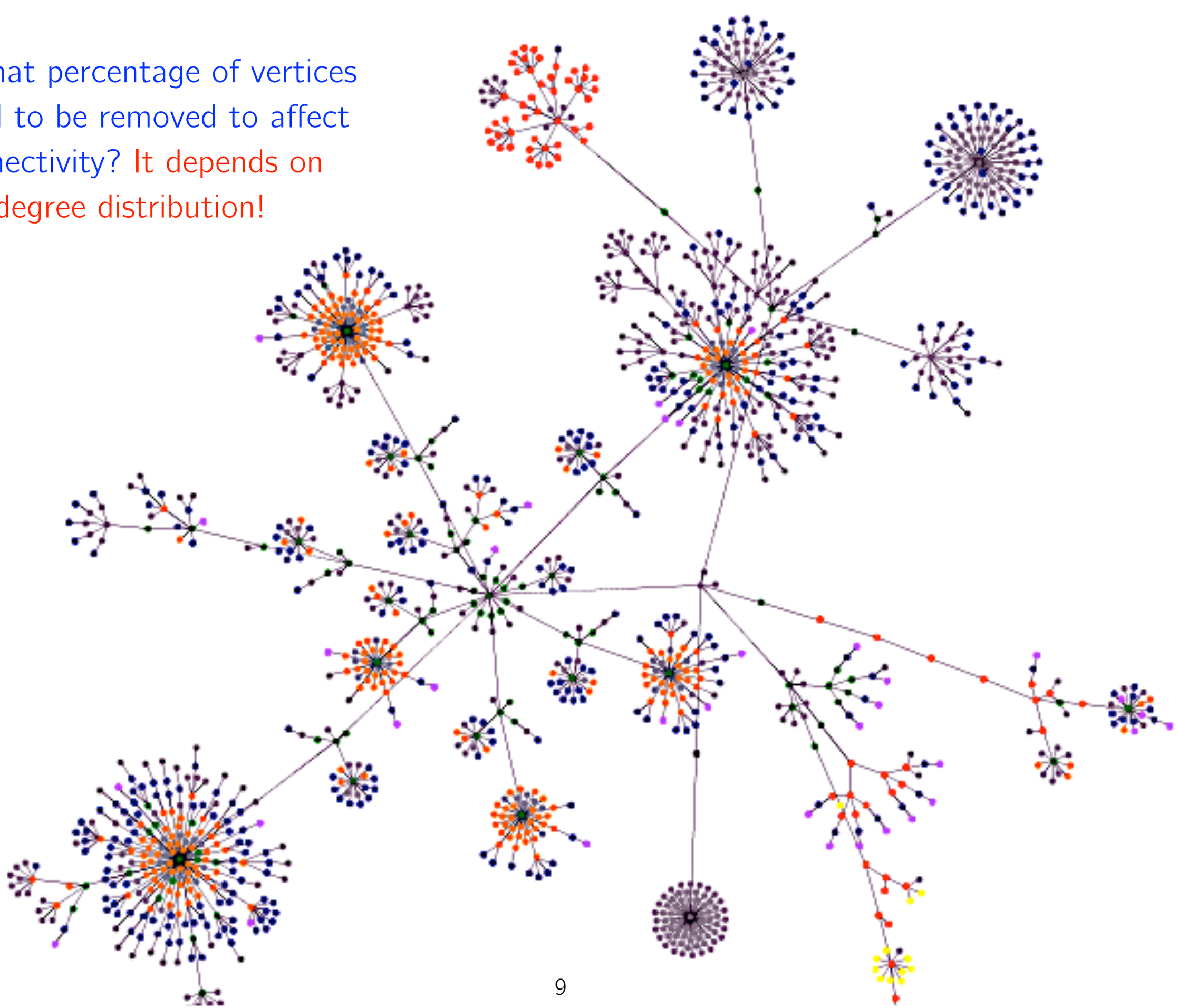


...which vertex would proof most crucial?

...what percentage of vertices need to be removed to affect connectivity?



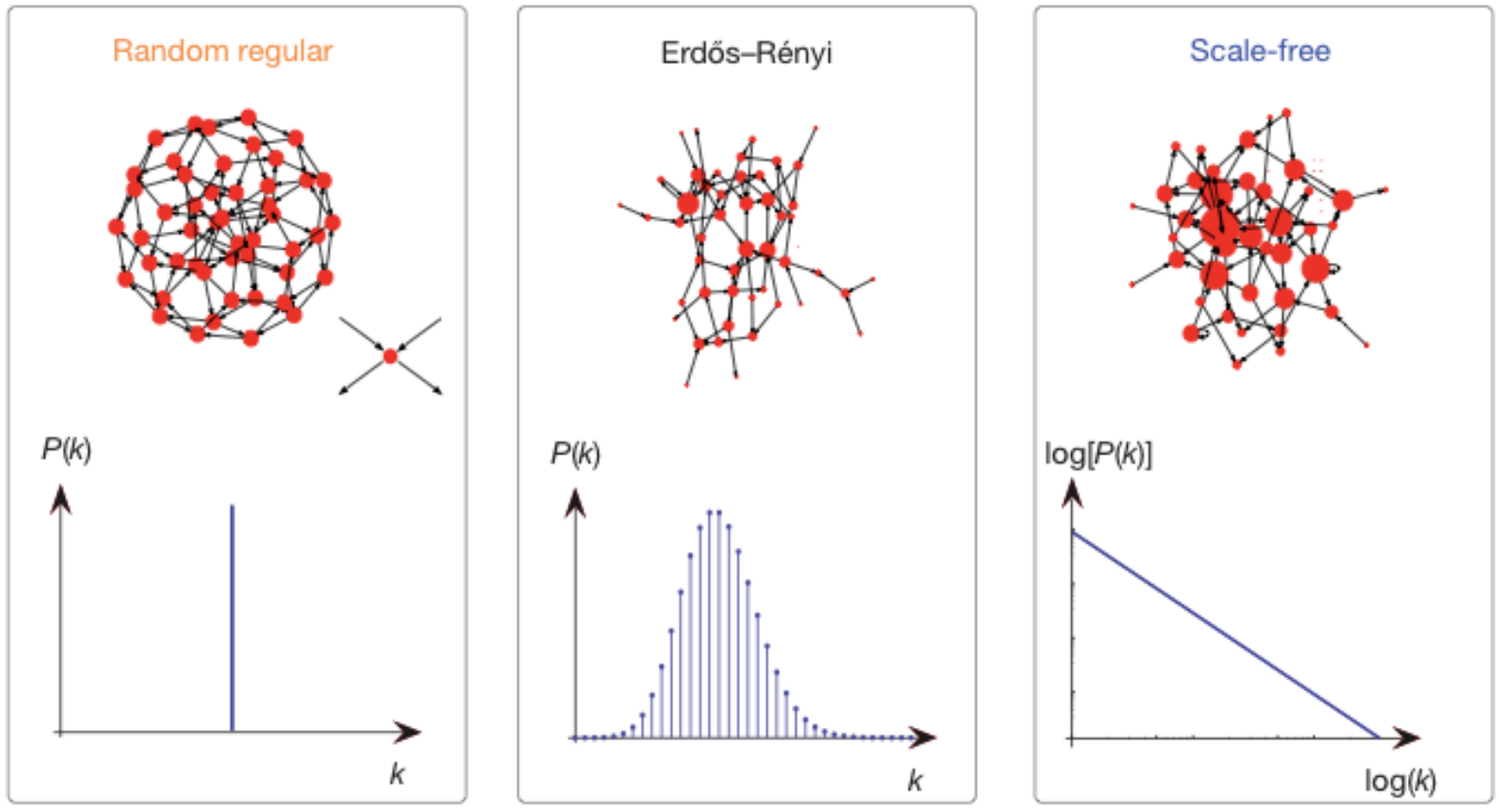
...what percentage of vertices need to be removed to affect connectivity? It depends on the degree distribution!



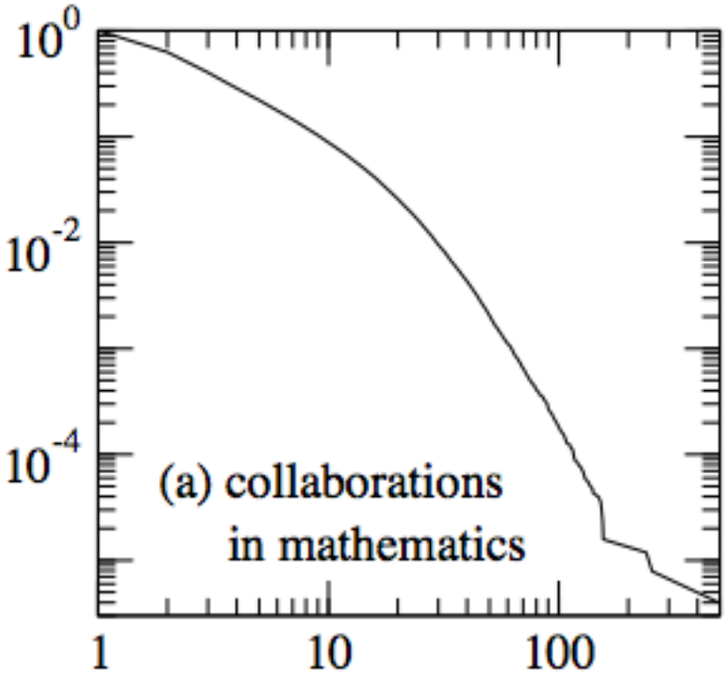
Review

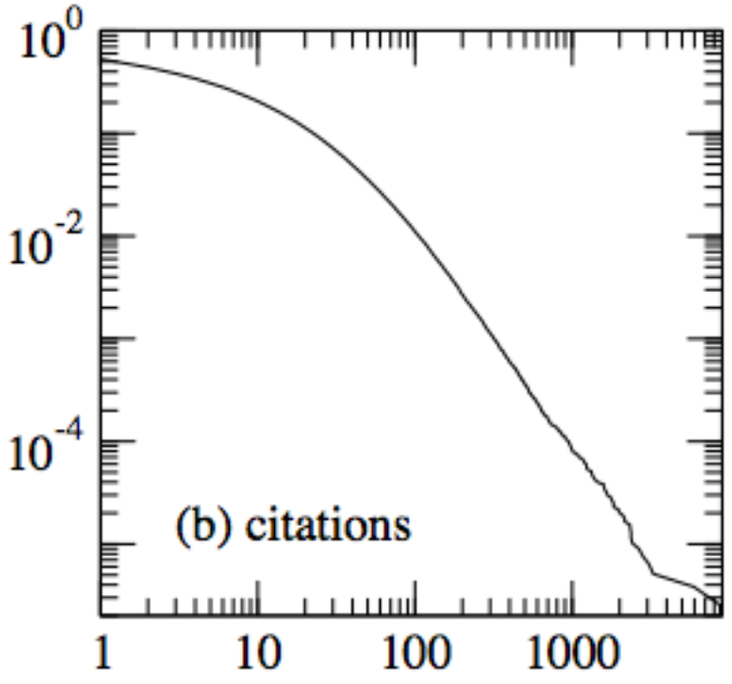
Common Properties (degree distributions):

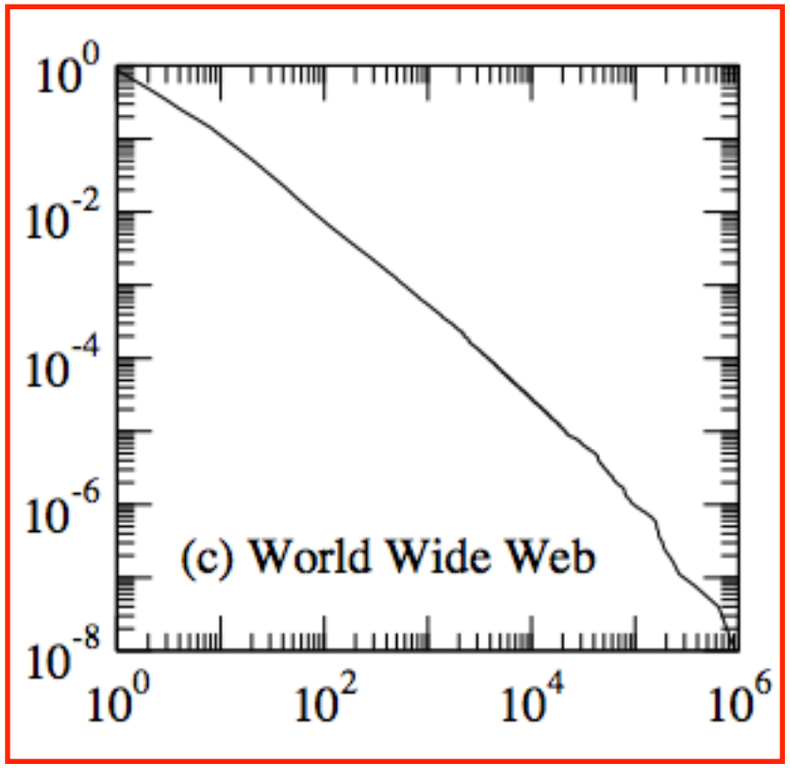
1. regular
2. random
3. heavy-tailed
   - power-law
   - exponential

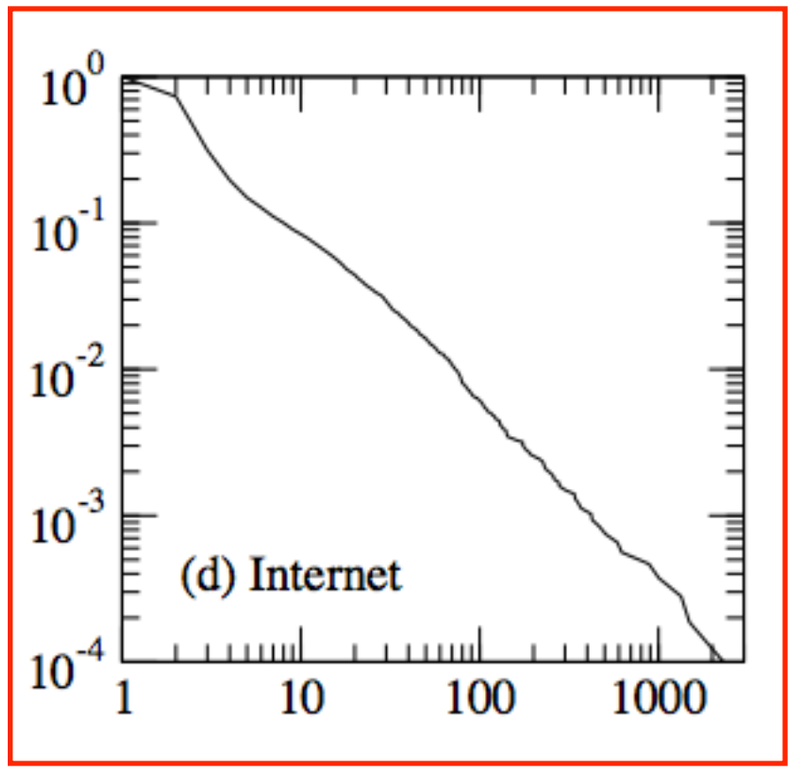


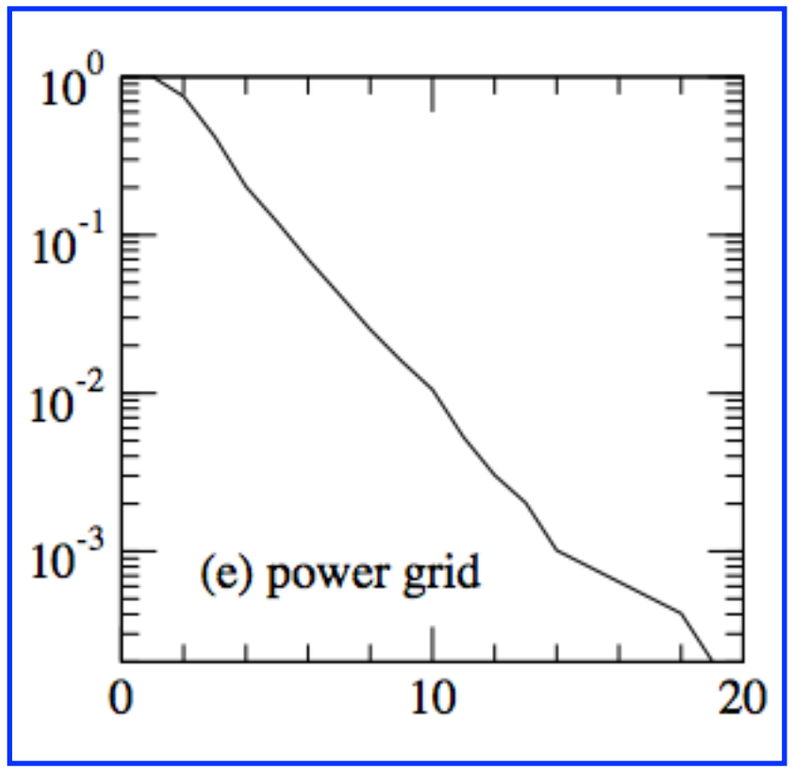
From Yang-Yu Lui et al., Controllability of complex networks

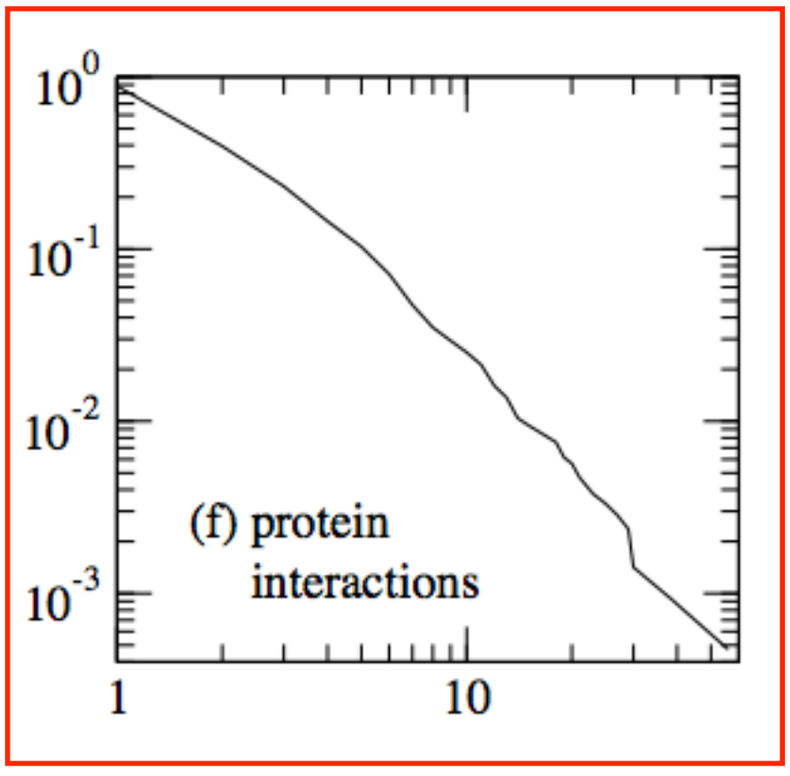

10^0
10^{-2}
10^{-4}
(a) collaborations
in mathematics
1
10
100


10^0
10^{-2}
10^{-4}
(b) citations
1
10
100
1000


10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
(c) World Wide Web
10^0
10^2
10^4
10^6

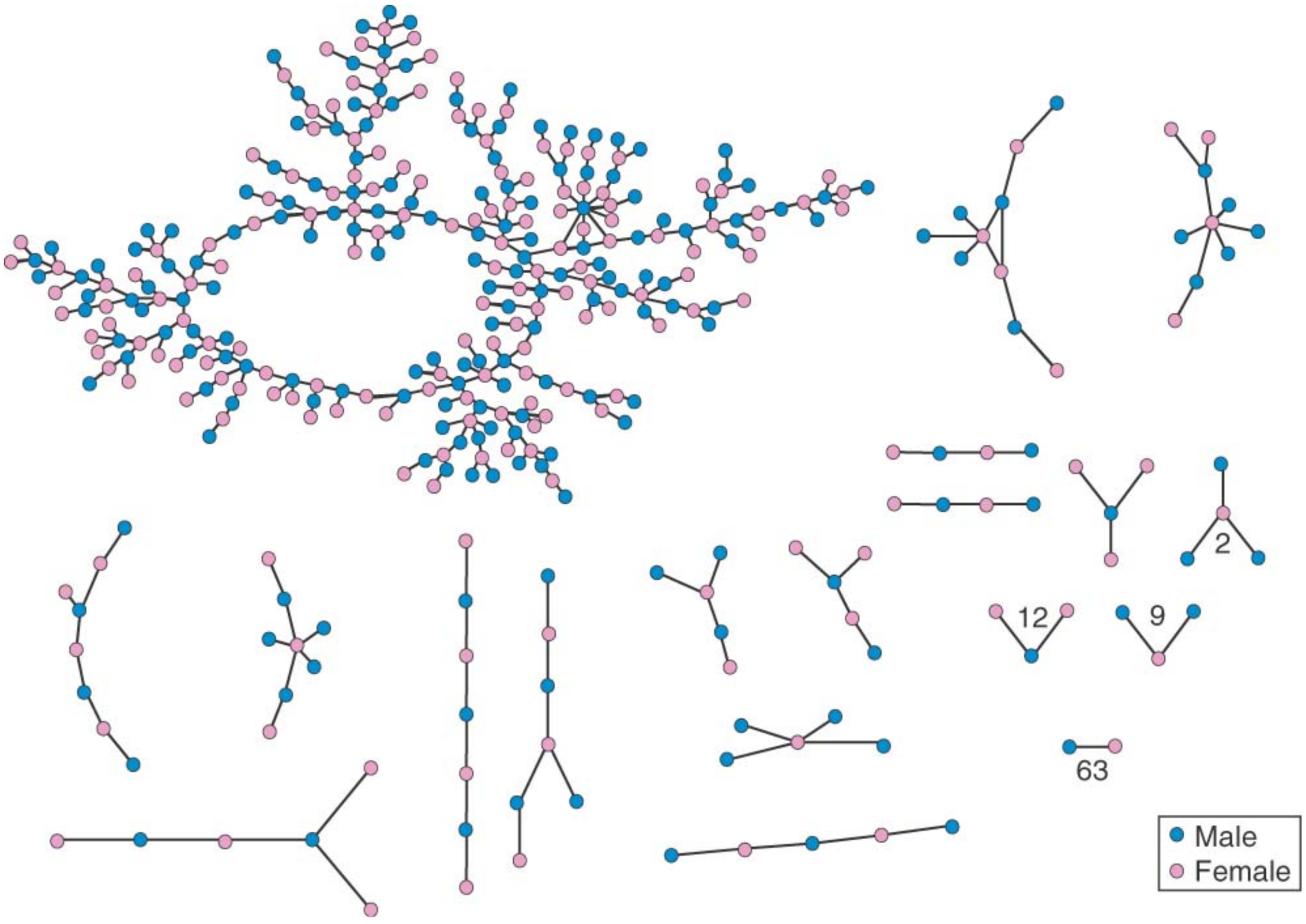

10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
(d) Internet
1
10
100
1000


10^0
10^{-1}
10^{-2}
10^{-3}
(e) power grid
0
10
20


10^0
10^{-1}
10^{-2}
10^{-3}
(f) protein
interactions
1
10

	network	type	number of nodes n	number of edges m	average degree z	average distance ℓ	scaling exponent α	global clustering $C^{(1)}$	av. local clustering $C^{(2)}$	deg. corr. r
social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208
	company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88	0.276
	math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34	0.120
	physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.45	0.56	0.363
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60	0.127
	telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1			
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16	
	email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13	0.092
	student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001	−0.029
	sexual contacts	undirected	2 810				3.2			
information	WWW `nd.edu`	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	−0.067
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7			
	citation network	directed	783 339	6 716 198	8.57		3.0/–			
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	–	0.13	0.15	0.157
	word co-occurrence	undirected	460 902	17 000 000	70.13		2.7		0.44	
technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39	−0.189
	power grid	undirected	4 941	6 594	2.67	18.99	–	0.10	0.080	−0.003
	train routes	undirected	587	19 603	66.79	2.16	–		0.69	−0.033
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082	−0.016
	software classes	directed	1 377	2 213	1.61	1.51	–	0.033	0.012	−0.119
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030	−0.154
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011	−0.366
biological	metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.090	0.67	−0.240
	protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.072	0.071	−0.156
	marine food web	directed	135	598	4.43	2.05	–	0.16	0.23	−0.263
	freshwater food web	directed	92	997	10.84	1.90	–	0.20	0.087	−0.326
	neural network	directed	307	2 359	7.68	3.97	–	0.18	0.28	−0.226

Relationship structure at a high school



P. S. Bearman et al., "Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks." American Journal of Sociology 110, 44-91 (2004)

Formation mechanism

- Different underlying mechanism
- Same outcome
- Growth + preferential attachment
- Growth + triad formation

Key questions (Homework 1)

- Can both models generate the scaling exponent associated to high school relationships?
- Can both models generate similar clustering coefficients?
- Are the average node-node distances similar?

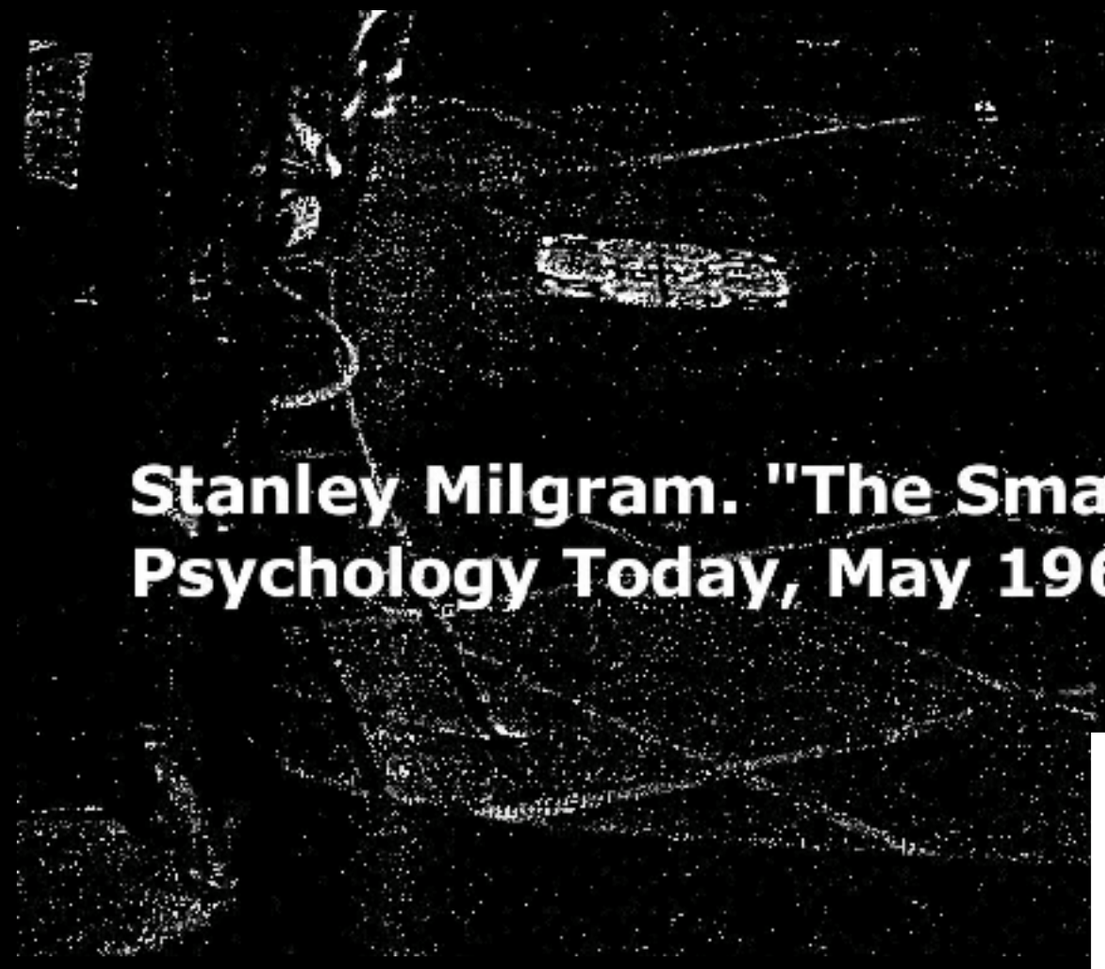
Today

- Small-world effect
- Transitivity or clustering
- Strong and weak ties

Later...

- Homophily
- Mechanisms underlying homophily
- Community structure

Small-word



Stanley Milgram. "The Small World Problem". Psychology Today, May 1967.

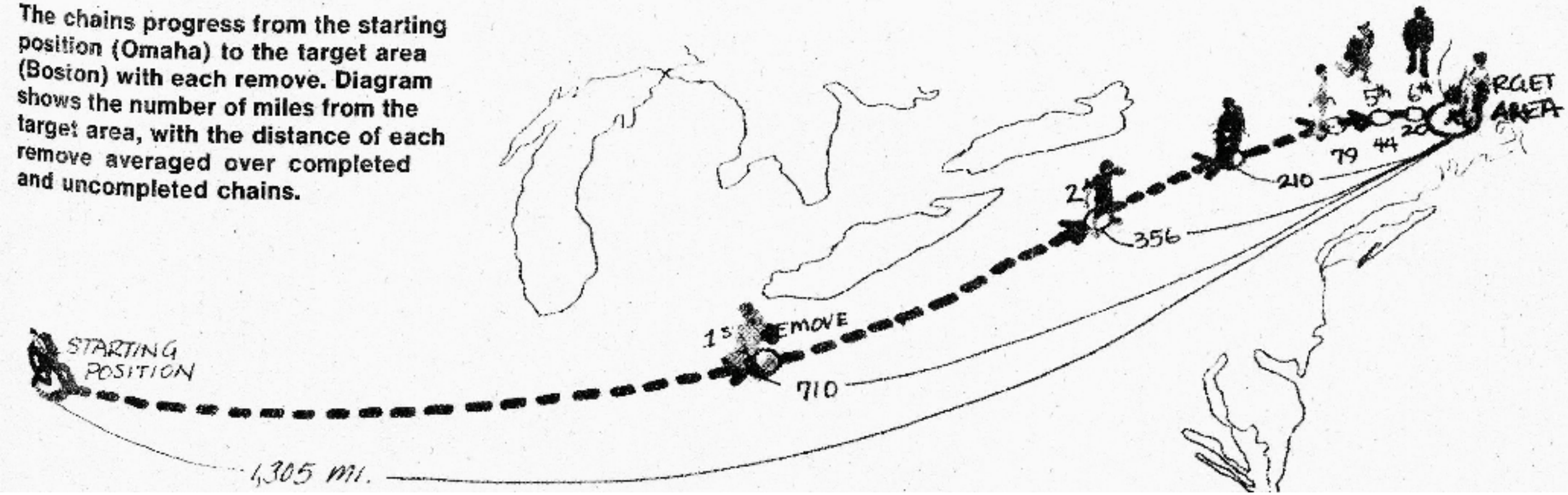
The Small-World

By Stanley Milgram

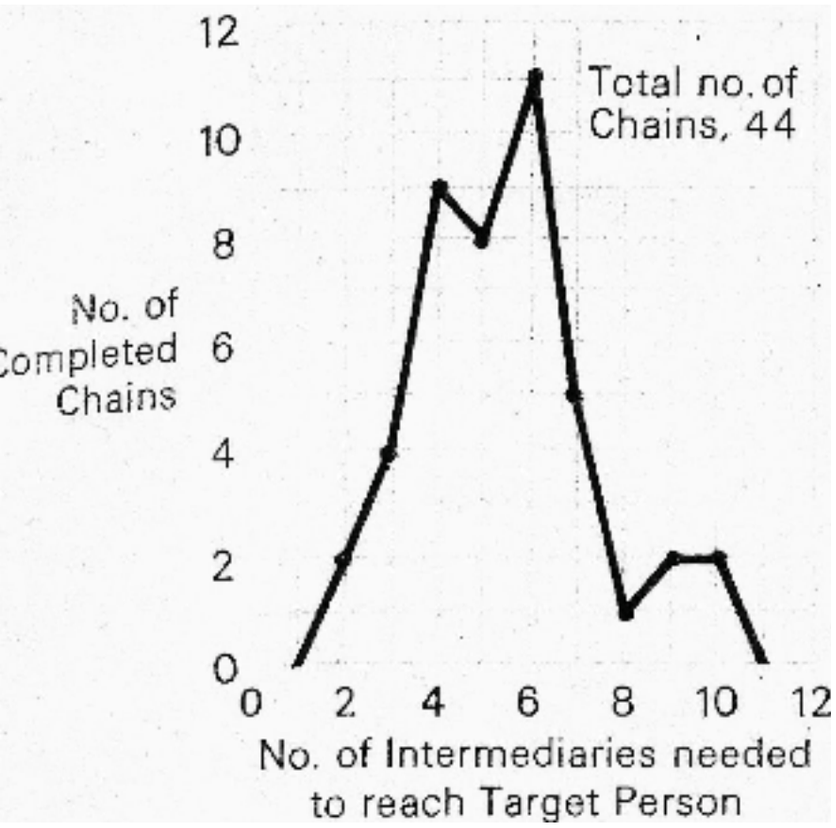
Almost all of us have had the experience of encountering someone far from home, who, to our surprise, turns out to share a mutual acquaintance with us. This kind of experience occurs with sufficient frequency so that our language even provides a cliché to be uttered at the appropriate moment of recognizing mutual acquaintances.
We say, "My it's a small world."

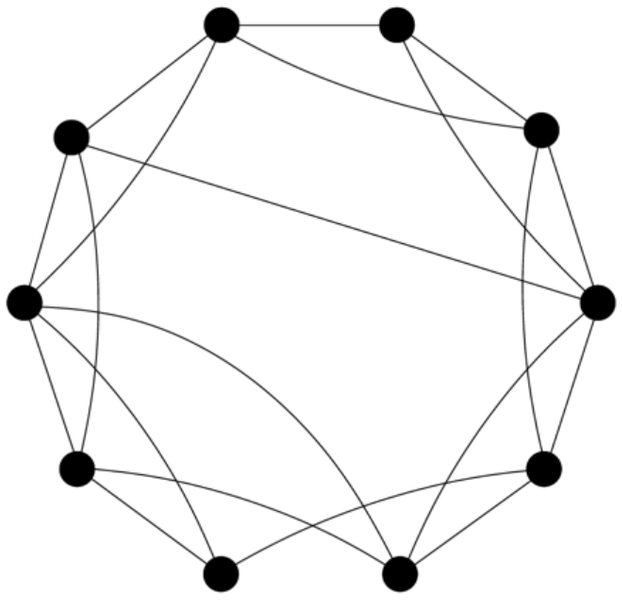
Assumptions

- The name of the target person
- Additional information (orient the participants towards specific individuals)
- If you do not know the target person on a first-hand bases, do not try to contact him directly
- A roster on which each person in the chain writes his name (prevent endless looping)



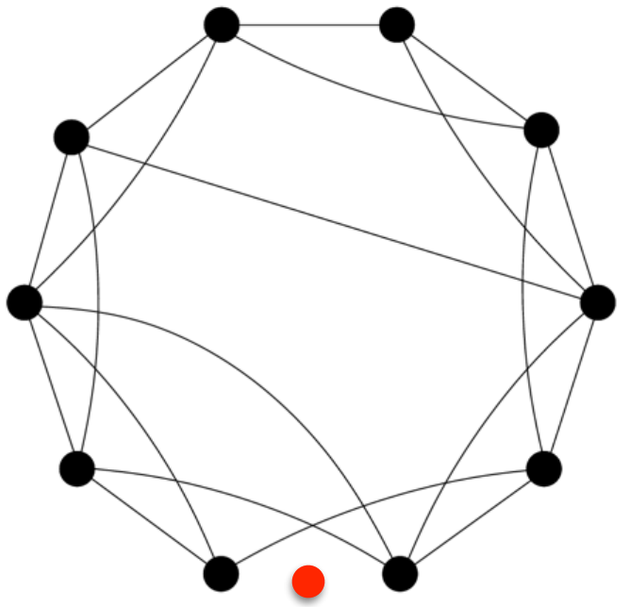
from Jeffrey Travers and Stanley Milgram, "An Experimental Study of the Small World Problem" Phycology Today, 1967





$$\ell = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}$$

total possible edges in an undirected network of n nodes

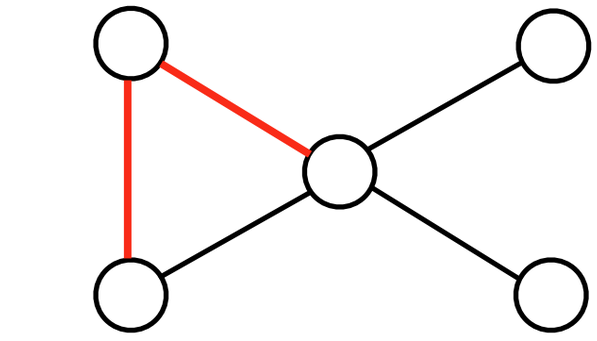


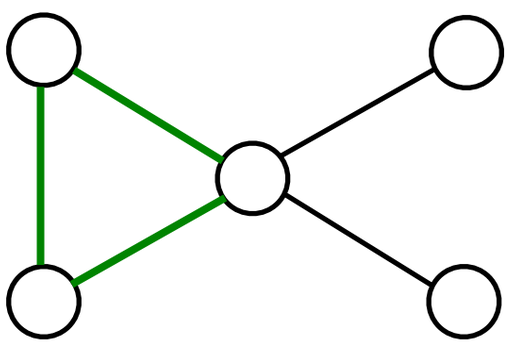
$$\ell^{-1} = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d_{ij}^{-1}$$

	network	type	number of nodes n	number of edges m	average degree z	average distance ℓ	scaling exponent α	global clustering $C^{(1)}$	av. local clustering $C^{(2)}$	deg. corr. r
social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208
	company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88	0.276
	math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34	0.120
	physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.45	0.56	0.363
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60	0.127
	telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1			
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16	
	email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13	0.092
	student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001	−0.029
	sexual contacts	undirected	2 810				3.2			
information	WWW `nd.edu`	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	−0.067
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7			
	citation network	directed	783 339	6 716 198	8.57		3.0/–			
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	–	0.13	0.15	0.157
	word co-occurrence	undirected	460 902	17 000 000	70.13		2.7		0.44	
technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39	−0.189
	power grid	undirected	4 941	6 594	2.67	18.99	–	0.10	0.080	−0.003
	train routes	undirected	587	19 603	66.79	2.16	–		0.69	−0.033
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082	−0.016
	software classes	directed	1 377	2 213	1.61	1.51	–	0.033	0.012	−0.119
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030	−0.154
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011	−0.366
biological	metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.090	0.67	−0.240
	protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.072	0.071	−0.156
	marine food web	directed	135	598	4.43	2.05	–	0.16	0.23	−0.263
	freshwater food web	directed	92	997	10.84	1.90	–	0.20	0.087	−0.326
	neural network	directed	307	2 359	7.68	3.97	–	0.18	0.28	−0.226

Transitivity

A friend of a friend is also my friend

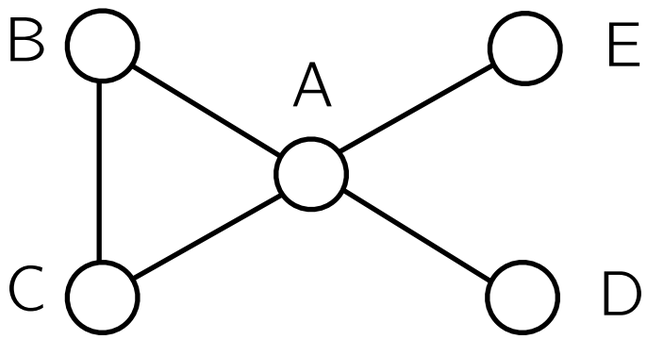




Transitivity coefficient

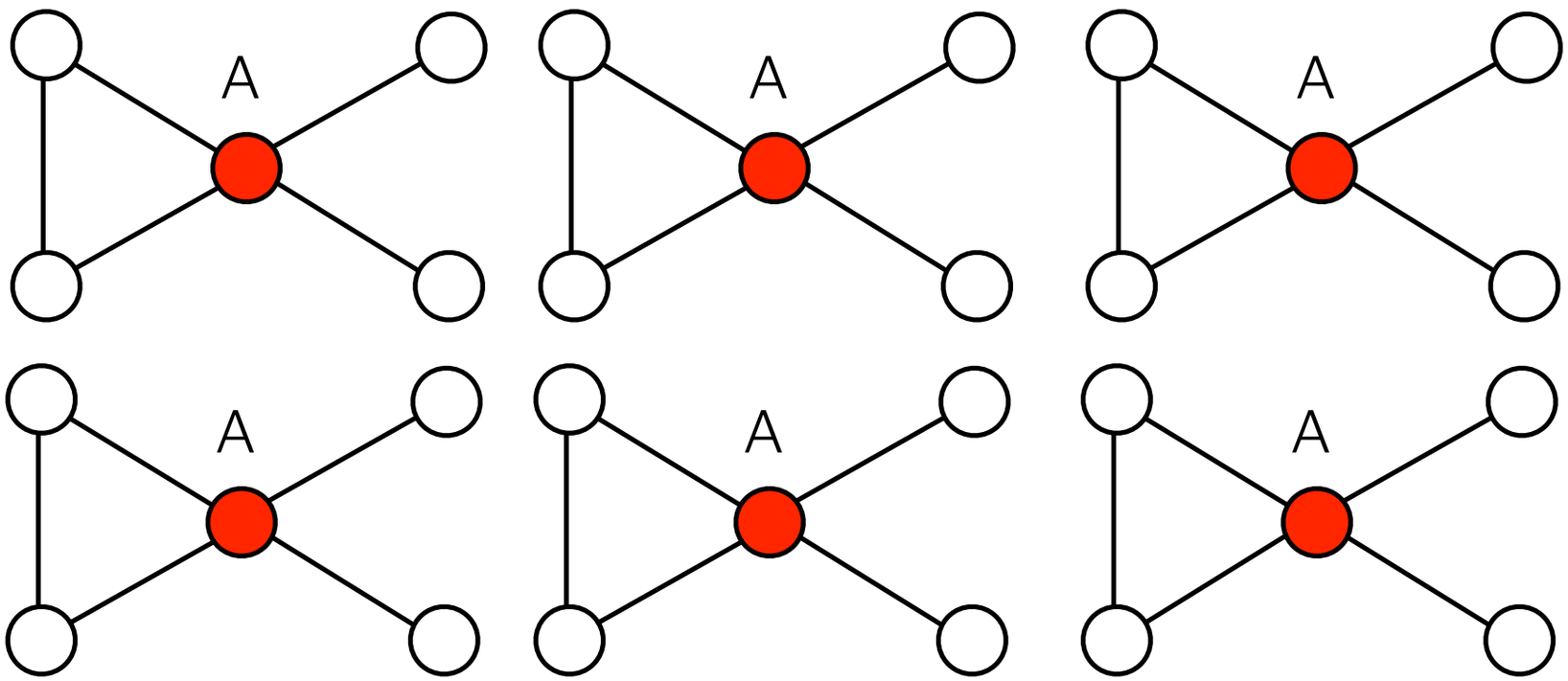
$$C = 3\frac{\text{number of triangles}}{\text{number of triples of vertices}} = \frac{3}{8}$$

"fraction of transitive triples"

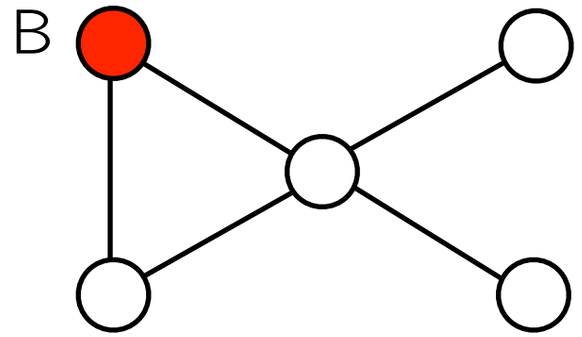


$$C_i = 3\frac{\text{number of triangles connected to vertex } i}{\text{number of triples centered on vertex } i}$$

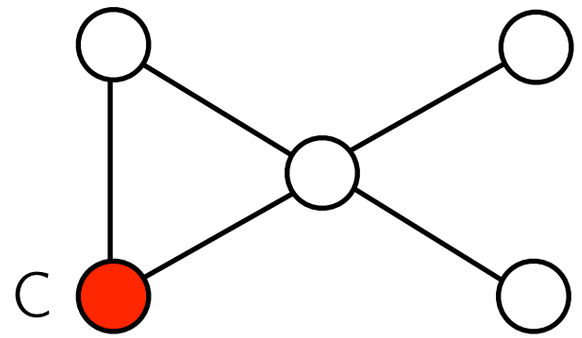
$$C = \frac{1}{n}\sum_i C_i = \frac{13}{30}$$



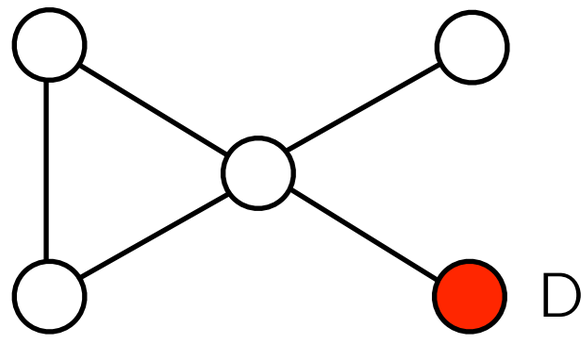
$$C_A = \frac{1}{6}$$



$C_B = 1$



$C_C = 1$



$C_C = C_E = 0$

Ratio of the means vs. mean of the ratios

Ratio of the means

$$C = 3\frac{\text{number of triangles}}{\text{number of triples of vertices}} = \frac{3}{8}$$

Mean of the ratio

$$C_i = 3\frac{\text{number of triangles connected to vertex } i}{\text{number of triples centered on vertex } i}$$

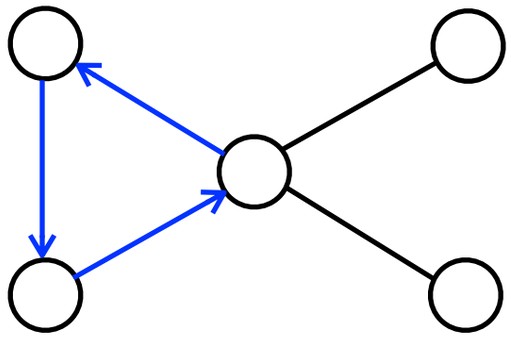
$$C = \frac{1}{n}\sum_i C_i = \frac{13}{30}$$

	network	type	number of nodes n	number of edges m	average degree z	average distance ℓ	scaling exponent α	global clustering $C^{(1)}$	av. local clustering $C^{(2)}$	deg. corr. r
social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208
	company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88	0.276
	math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34	0.120
	physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.45	0.56	0.363
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60	0.127
	telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1			
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16	
	email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13	0.092
	student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001	−0.029
	sexual contacts	undirected	2 810				3.2			
information	WWW `nd.edu`	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	−0.067
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7			
	citation network	directed	783 339	6 716 198	8.57		3.0/–			
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	–	0.13	0.15	0.157
	word co-occurrence	undirected	460 902	17 000 000	70.13		2.7		0.44	
technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39	−0.189
	power grid	undirected	4 941	6 594	2.67	18.99	–	0.10	0.080	−0.003
	train routes	undirected	587	19 603	66.79	2.16	–		0.69	−0.033
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082	−0.016
	software classes	directed	1 377	2 213	1.61	1.51	–	0.033	0.012	−0.119
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030	−0.154
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011	−0.366
biological	metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.090	0.67	−0.240
	protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.072	0.071	−0.156
	marine food web	directed	135	598	4.43	2.05	–	0.16	0.23	−0.263
	freshwater food web	directed	92	997	10.84	1.90	–	0.20	0.087	−0.326
	neural network	directed	307	2 359	7.68	3.97	–	0.18	0.28	−0.226

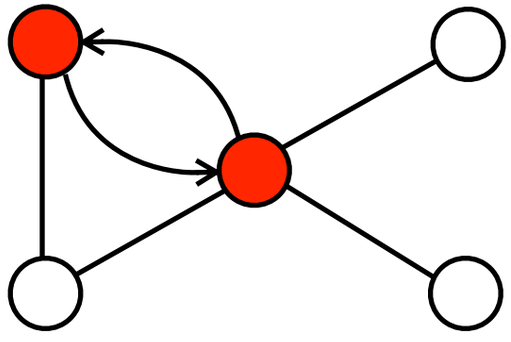
Remarks

- Considerably higher than for random graphs
- Measures density of triangles
- Generalize to four and longer loops
- Important in directed graphs

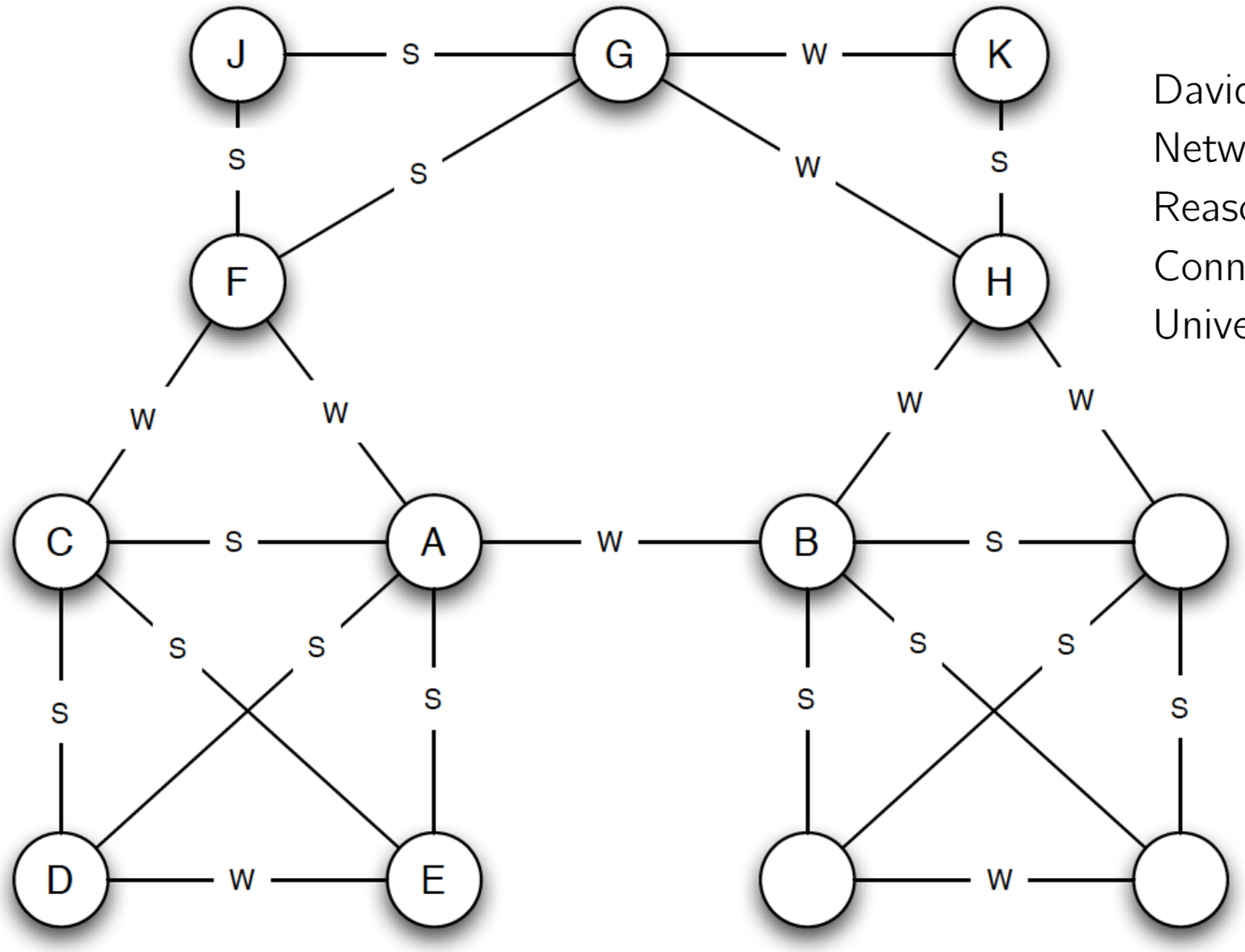
feedback cycle of length 3



reciprocity



Weak and strong links



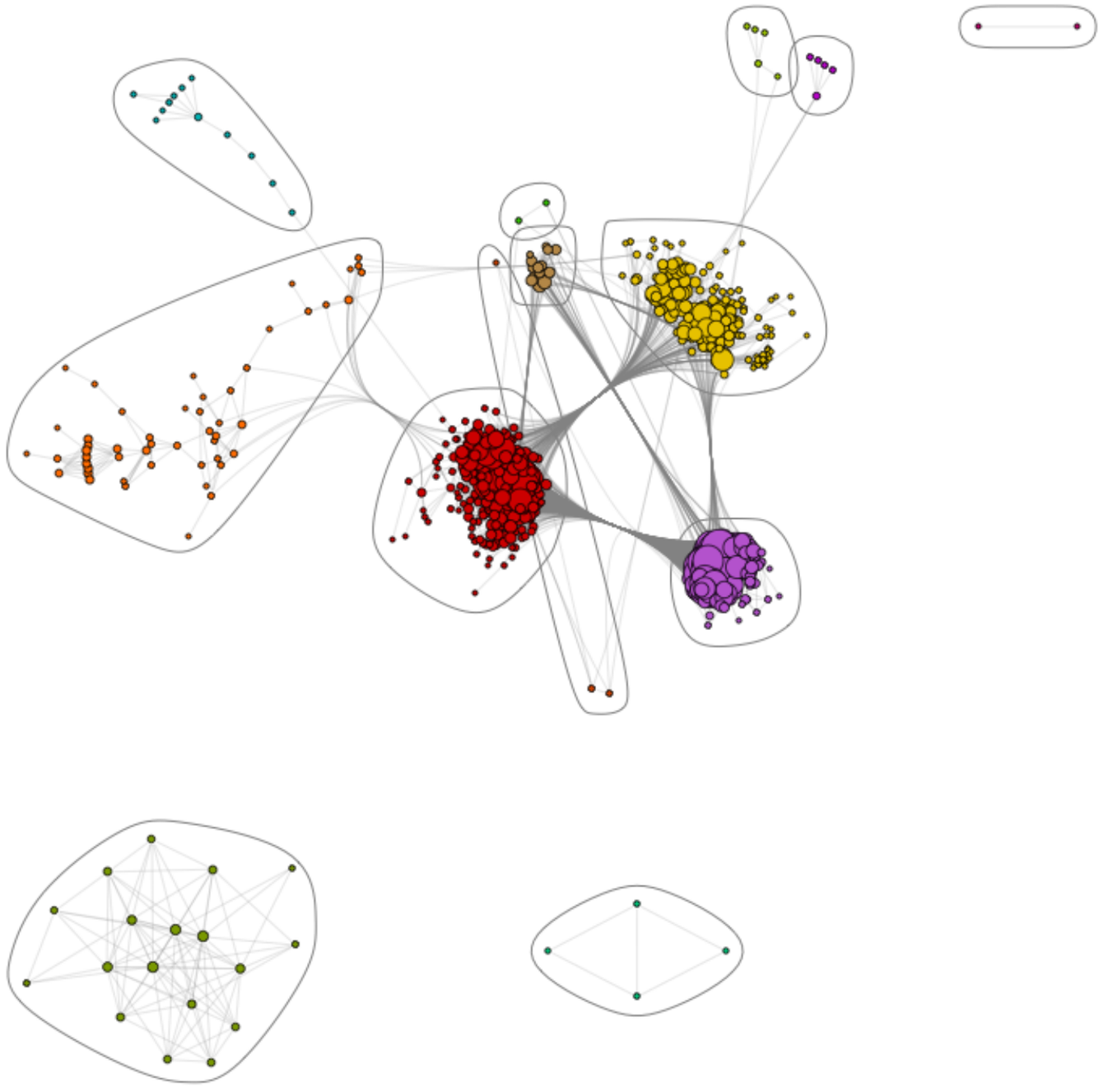
David Easley and Jon Kleinberg. Networks, Crowds, and Markets: Reasoning about a Highly Connected World. Cambridge University Press, 2010.

Strong Triadic Closure

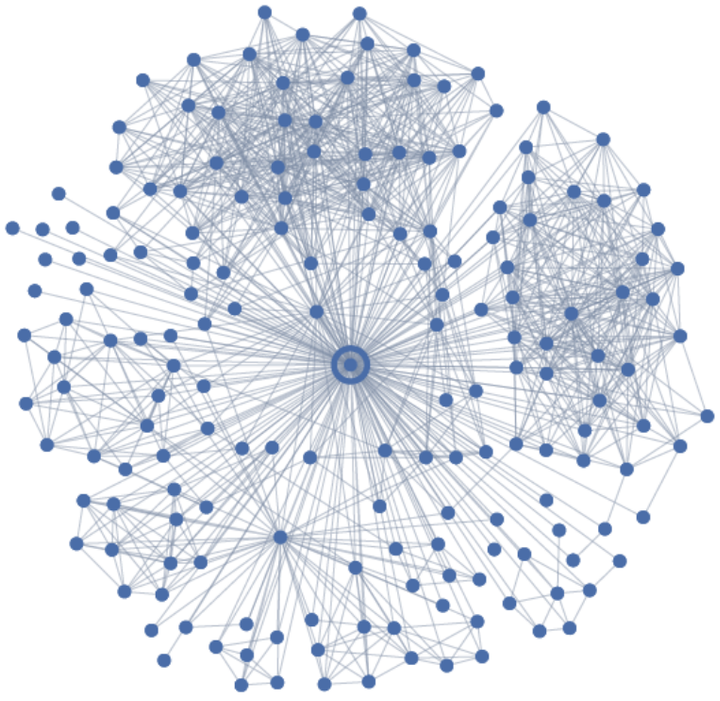
We say that a node A violates the Strong Triadic Closure Property if it has strong ties to two other nodes B and C, and there is no edge at all (either a strong or weak tie) between B and C. We say that a node A satisfies the Strong Triadic Closure Property if it does not violate it.

Claim

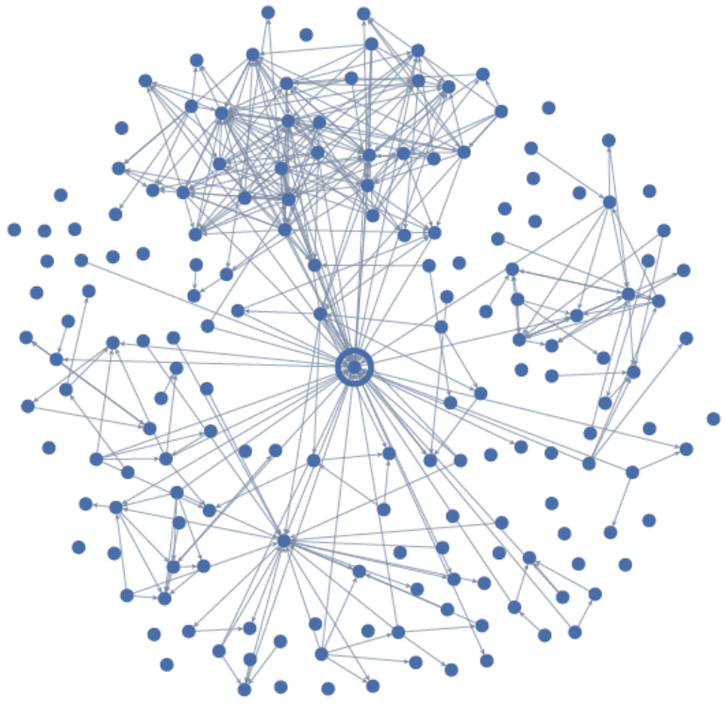
If a node A in a network satisfies the Strong Triadic Closure Property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie.



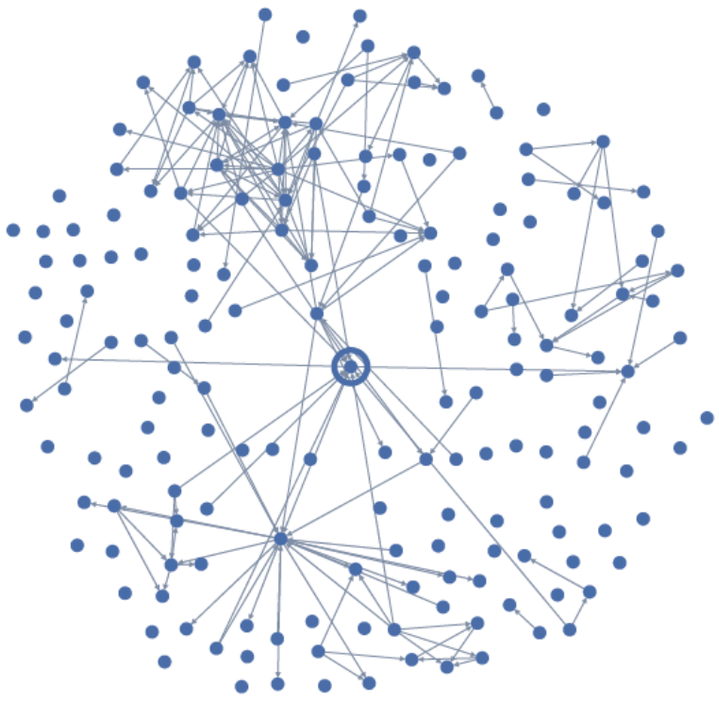
All Friends



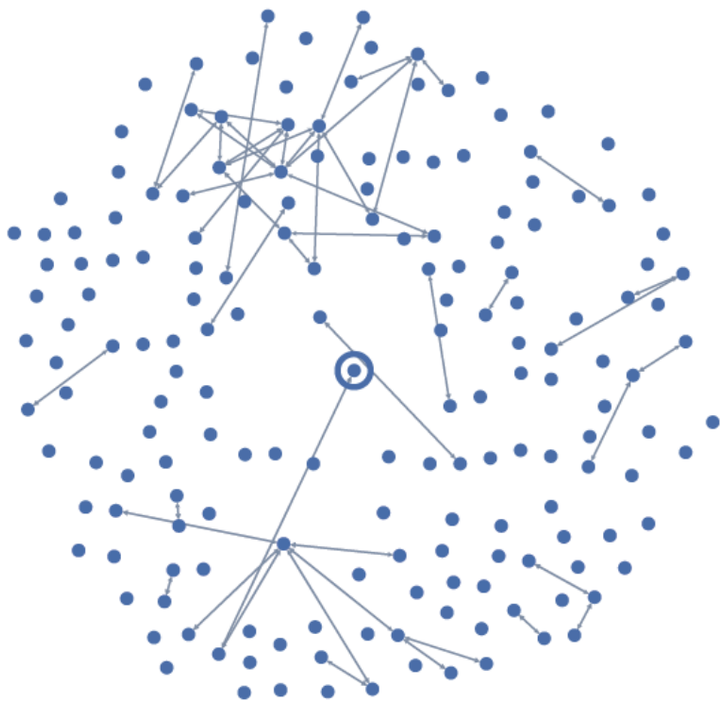
Maintained Relationships



One-way Communication



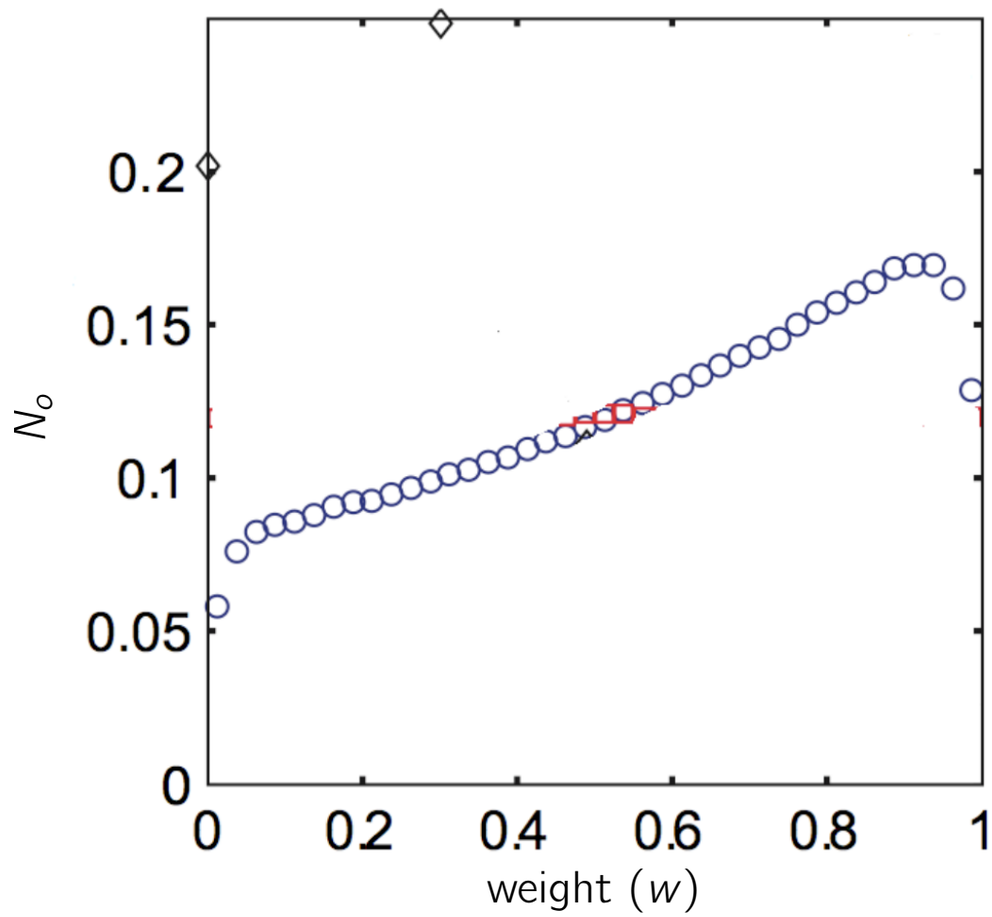
Mutual Communication



Cameron Marlow, Lee Byron, Tom Lento, and Itamar Rosenn. Maintained relationships on Facebook, 2009. On-line at http://overstated.net/2009/03/09/maintainedrelationships-on-facebook.

Neighborhood overlap

$$N_o = \frac{\text{number of nodes that are neighbors of both A and B}}{\text{number of nodes that are neighbors of at least A or B}}$$



Claim

If a node A in a network satisfies the Strong Triadic Closure Property and is involved in at least two strong ties, then any local bridge it is involved in must be a weak tie.

J.-P. Onnela, et al. Structure and tie strengths in mobile communication networks. Proc. Natl. Acad. Sci. USA, 104:7332-7336, 2007.

Next class

- Homophily
- Mechanisms underlying homophily
- Community structure