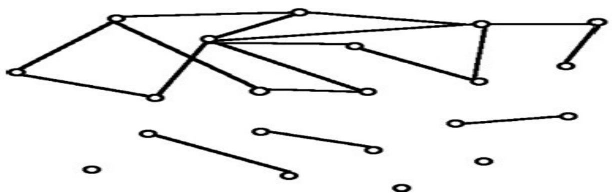
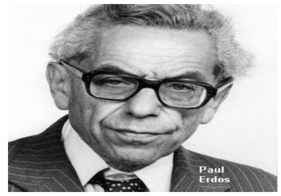
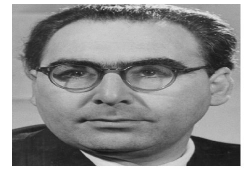
Lecture 10: Random Graphs, Small World Networks, and Power Laws in Networks-- Feb 27

Paul Erdos and Alfred Renyi sat in a Budapest cafe in late 1950s and put dots on paper and pondered # of links that would **connect them** if random linked with probability p → GRAPH THEORY







Terminolology: G(n,p). n = #vertices of graph; p is probability one vertex links to another. There are n(n-1)/2 ~$n^2/2$ possible edges, where /2 reflects the fact that each edge links 2 vertexes .

Vertices	possible edges
n	n(n-1)/2
3	3
4	6
10	45
16	120
250k	32 billion

Actual graphs have many fewer connections/edges than the max.
Edges are often clustered so that if one vertex connects to two others, the two others have higher probability of being linked than p – ie your two friends probably know each other
A key concept is connectedness – the set of vertices among whom it is possible to get from one to any other through some path – the network. Connected graphs have small diameter, where diameter is shortest path between any two vertices.

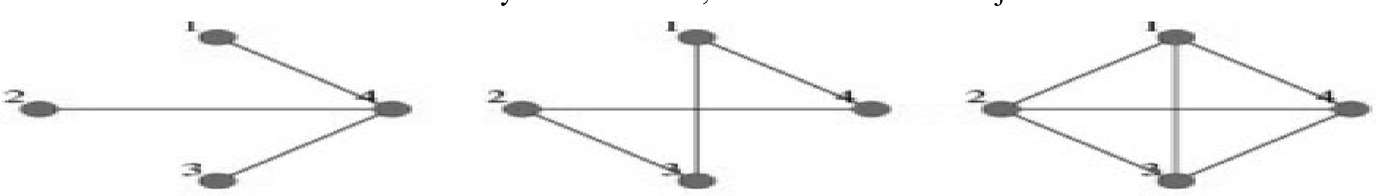
In people networks, vertices are us while the edges show interactions**. Directional graph** distinguishes A –> B from B–> A. **Web is giant digraph**, with pages as vertices and edges as connections directing you from one site to another.

Erdos and Renyi published paper in 1959 in Publicationes Mathematicae Debrecen founded by A. Renyi, T. Szele and O. Varga in 1950. Institute of Mathematics, University of Debrecen, Hungary. For 40 -45 years little attention to paper with just a few citations, Then # of citations explodes in 2000s so that in 2017 it had 18027 cites

Definitions: V(G) the # n of vertices is called **ORDER OF A GRAPH:** 10 vertices –> graph has order 10
E(G) is edge list – M edges gives size of graph. Since max edges is n(n-1)/2 graph of order 10 has 45 edges at max
Degree k_v of vertex is the # of edges to which it connects. **Average degree** of graph k = M/n a graph with 16 edges and 8 vertices, you have average degree of 2. **Regular graph** all vertices with same degree. Graphs have natural **distance metric d(i,j)** defined as minimum # of edges you need to get from i to j. **Diameter** of graph is Max(i,j) d(i,j) – the biggest distance between any two vertices.

Characteristic path length L(G) measures average distance btwn vertices = Average of d(i,j) over all i,j – if you have distances of 5, 7, 6, you have characteristic path length of 6.

Cluster coefficient measures how closely linked –) a highly clustered neighborhood. For v C(v) defined as fraction of possible directed edges among k neighbors that exist (k) = k!/[(k-2)! 2!] = k (k-1)/2. Take a vertex, look at all its neighbors, find links and divide actual links by max. **Cluster coefficient C** is average C(v). Graphs represented by **adjacency matrix** with rows and columns labeled by vertices and 1,0 to reflect whether adjacent or not.



$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Now think about Coronavirus possible pandemic. You sit in cafe in Cambridge next to someone who just came from Paris who shook hands with someone in Korea who ...

Social network. In 1967 Stanley Milgram asked people in Omaha to pass letters to unknown persons in Mass through someone they knew. The people who agreed and reached someone did so in about 6 links ... Six degrees of separation, which seemed surprisingly small. With Internet and Facebook and LinkedIn would presumably be faster to send a message to some person with only their name and city/state as identifying information through link of known people: you know someone at Harvard from Nebraska or who played against Neb in sport or … who knows politician –

In graph theory model the nodes are people, the links are the letter. But nodes can be scientists and the edge could be writing a paper together – the coauthor network. Or nodes could be the paper and the edge whether paper cited another paper – citation network. Or nodes could be actors and edges could be movies.

Watts "replication of Milgrim: Duncan Watts (now at Microsoft) and group ((Science, Aug 8, 2003) Internet-based experiment in which participants were randomly allocated one of 18 target persons from 13 countries and told to find them. **DID NOT FIND MANY CONNECTS**. Completed degrees of separation around 4, but ....

"The presence of highly connected individuals (hubs) appears to have limited relevance to the kind of social search embodied by our experiment. But only 384 of 24,163 chains reached their targets. Chains may have terminated randomly, because of individual apathy or disinclination to participate. OR because they would never get to final person. If people searching for remote targets do not have sufficient incentives to proceed, the small-world hypothesis will not appear to hold... empirically observed network structure can only be meaningfully interpreted in light of the actions, strategies, and even perceptions of the individuals embedded in the network: **Network structure alone is not enough for people to connec6. NEED economic incentives to motivate people to keep pushing...**

In academic research network – can link through co-authors or through citations. If you want to connect to a chemist in Germany what would you do? Academics have dense local networks but links to outside researchers. Field is more important than geography in making quick links.

TWO MAIN AREAS BEYOND RANDOM MODEL: 1)–SMALL WORLD – Length and Connectedness

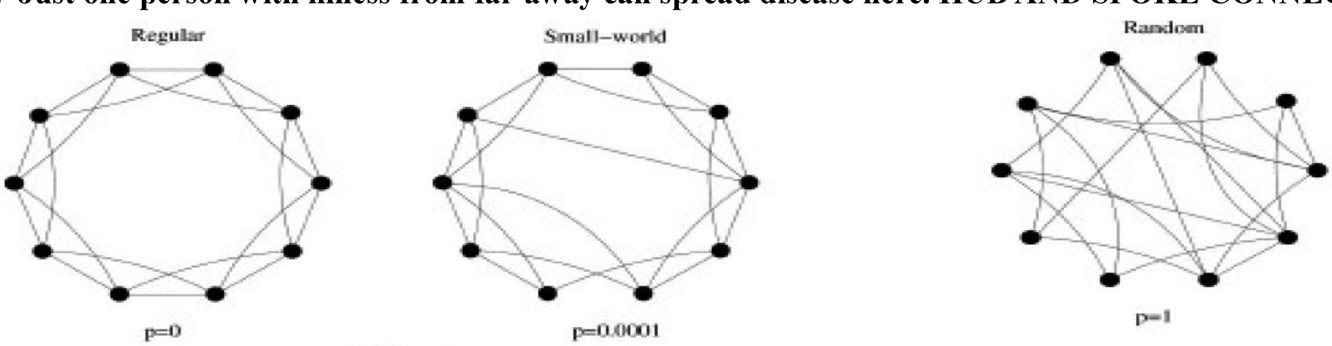
Random graph –You are randomly connected to anyone regardless of geography. Key results: Many graph properties, like connectiveness measure of all vertexes reachable from any other-- suddenly appear at a threshold level, where p ~ 1/n. Probability that a vertex has k edges is a Poisson distribution. T**heorem:** As p increases from ~0 to 1/n connectedness gets huge, with threshold for connectedness of $G(n, p)$ of (ln n)/ n. If $p < 1/n$, graph in $G(n, p)$ will almost surely have no connected components > than $O(\log n)$

9/ 11 example of random graph. How "close" people are to someone 3,000 killed in the terrorist attack

\# who know people in event = (# people in personal network) x (# people in event) = 300 people in personal networks x 3,000 killed = 900,000. See http://nersp.osg.ufl.edu/~ufruss/documents/Estimating.network.size.pdf; and http://www.stat.washington.edu/~tylermc/pubs/mccormick_degree.pdf for estimates of personal networks. Since the 900,000 know 300 others –> 270,000,000 were 2 people away from someone killed in that event –whole country.

Think of regular graph as mailman who goes house to house … or in ancient days, door to door salesperson who would have a "territory of blocks" to sell vacuum cleaners or snake oil or girl scout/boy scout cookies. If you are going to spread disease, it can be contained within local nhood.

Small World Graph – Start with a graph where every vertex is connected to next door neighbors but where one edge is detached with probability P from neighbor and randomly rewired to some other edge/far-away link. P tunes the system (recall Kaufman parameter K which tuned the NK landscape model). The SW graph has connectedness close to the regular graph but path length comparable to the random graph. **All we need is one person in n-hood to link us to world. Just one person with illness from far-away can spread disease here. HUB AND SPOKE CONNECTION.**



Thus small-world lies between highly ordered regular and highly disordered random.

Regular: No one has long connections P=0, so characteristic path length is BIG. People on your block know only people next door. The characteristic path length grows linearly with n.

Random: P~1 characteristic path length will be around 3, which is small. C is just 3/1000. As you move on random graph along one edge you reach k new places. If you move 2 edges from the start you should reach on average k^2 new places. If the graph is fully connected, you would have to go $K^L = n$ steps to reach any vertex so that L=ln n/ ln K. There is a logarithmic increase in L with log increases in k

Small World: (but with "shortcut long links") Say P = .005. We connect ½ percent of people randomly. Get small characteristic path length, but large connectedness. Ln increase in L with increases in ln K.

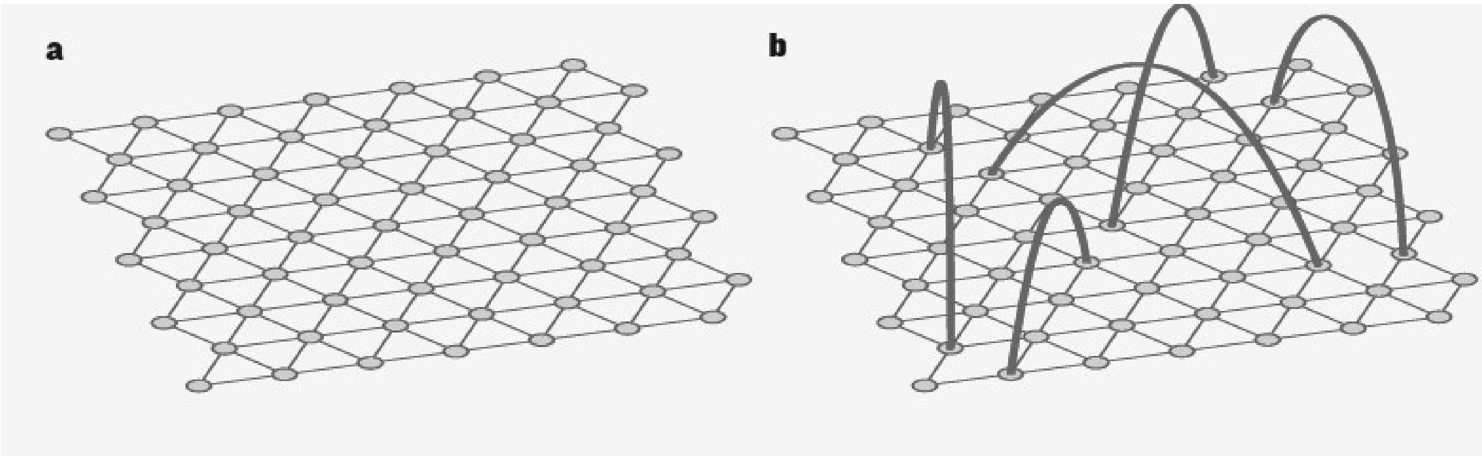


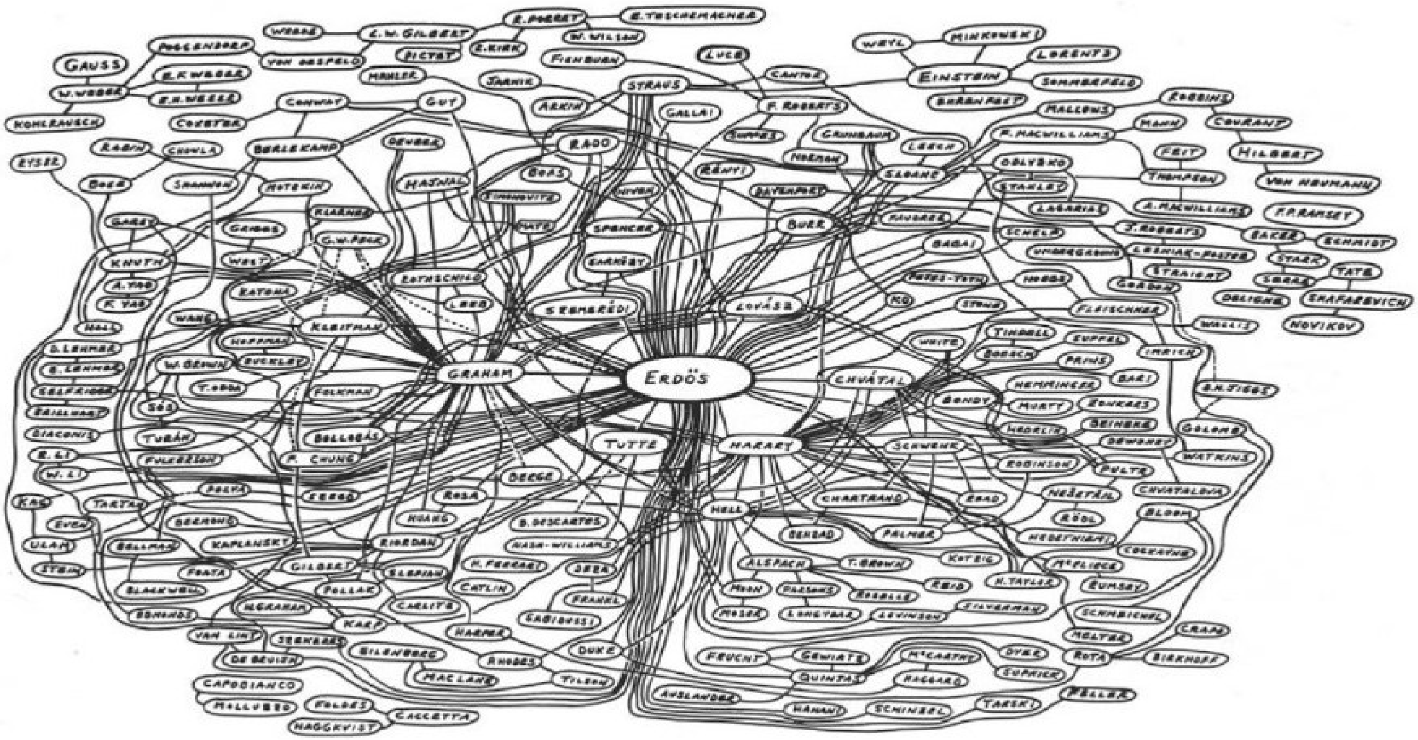
Figure 1 | The small-world network model. In 1998, Watts and Strogatz[1] described a model that helps to explain the structures of networks in the real world. **a**, They started with a regular network, depicted here as nodes connected in a triangular lattice in which each node is connected to six other nodes. **b**, They then allowed links between nodes to be rewired at random, with a fixed probability of rewiring for all links. As the probability increases, an increasing number of short cuts (red lines) connect distant nodes in the network. This generates the small-world effect: all nodes in the network can be connected by passing along a small number of links between nodes, but neighbouring nodes are connected to one another, forming clustered cliques. (Adapted from Samay/Vespignani.)

		L–Characteristic	Path Length	Connectedness	
Regular Graph	P~0,	~n/2k	(Large)	~3/4	(Large)
Small world	.001<P<.01	ln n/ln k	**(Small)**	**~3/4**	**(Large)**
Random graph	P~1	ln n/ln k	(Small)	~k/n	(Small)

The surprise is that o**nly a few shortcuts are needed to get across the whole space.** This contrasts with random where connection is low in neighborhood because neighbors could be far away. **What does this say to you about chances of getting coronavirus?**

A famous graph is the **Kevin Bacon Graph** … http://oracleofbacon.org/ which builds a map of actors and movies which answers three requests: <u>Find the link from Actor A to Actor B.</u> <u>How good a "center" is a given actor?</u> (Guess what a center is?) <u>Who are all the people with an Actor A number of N?</u> Consider the bacon numbers: Pinky Lee 1950s comedian) has a bacon number 2. Pinky was in Ocean's Eleven (1960) with Jessica James who was in Diner (1982) with Bacon. Andre the Giant has number 2. Charlie Chaplin has bacon number 3. Brigitte Bardot has bacon number 3.

An even more famous social graph – at least among mathematicians – is **Erdos graph** connects co-authors to Erdos (**Erdos number** 1), to people who wrote with a co-author (Erdos number 2) and so one. Most living mathematicians have finite Erdos number. Graph shows Erdos at center and links persons in arrays around him.



REAL PEOPLE DO NOT RANDOMLY CONNECT

Milgram type experiments depend on conscious choices of people. People "network" for conscious reasons. **Networking in job search**. Granoveter claimed that weak ties – people outside your immediate circle more likely to help you get a good job. Linked in makes its money by linking you to all kinds of people. Right Management, an arm of giant Manpower Group, analyzes data from 59,133 clients it advised over the last three years and finds:

Source of New Job

	2010	2009	2008
Networking	41%	45%	41%
Internet Job Board	25%	19%	19%
Agency/Search firm	11%	9%	11%
Direct Approach	8%	8%	8%
Online Network (2010)	4%	na	na
Advertisement	2%	7%	7%
Other	10%	12%	

So how do we model the use of LOCAL information to get close to the best short chain to connect? Intelligent choice should beat random probabilistic model.

Kleinberg sets up the problem on a two-dimensional grid, where $d(a,b; x,y) = |a-x| + |b-y|$ and postulates that **the chance that you link to far-away site v is proportionate to $d^{-r}(v)$ where r is the parameter for the model $(d^{-r}/\sum d^{-r}(v))$** The probability of getting a long link is proportional to d so for small r, say 0, equal chance of getting to any far away site. r = 0 implies uniform distribution – constant –> there exist "short" paths but a decentralized search cannot find them. You do not know where they are, so average links are going to be much greater than "smart search" $\sim n^{2/3} > \log(n)$. Say n is 100, then 22 steps instead of 7. If n is 1000, takes 100 steps instead of 10

With r=2 and the probability of getting a long link proportional to d^{-2} there is less chance of getting a long link, but going in right way helps. You do not know whom your contacts know but you know where the target is and that

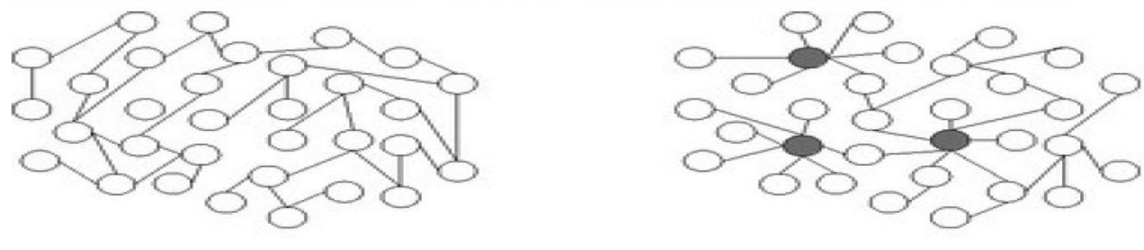
Table 1. Type, origin, and strength of social ties used to direct messages. Only the top five categories in the first two columns have been listed. The most useful category of social tie is medium-strength friendships that originate in the workplace.

Type of relationship	%	Origin of relationship	%	Strength of relationship	%
Friend	67	Work	25	Extremely close	18
Relatives	10	School/university	22	Very close	23
Co-worker	9	Family/relation	19	Fairly close	33
Sibling	5	Mutual friend	9	Casual	22
Significant other	3	Internet	6	Not close	4

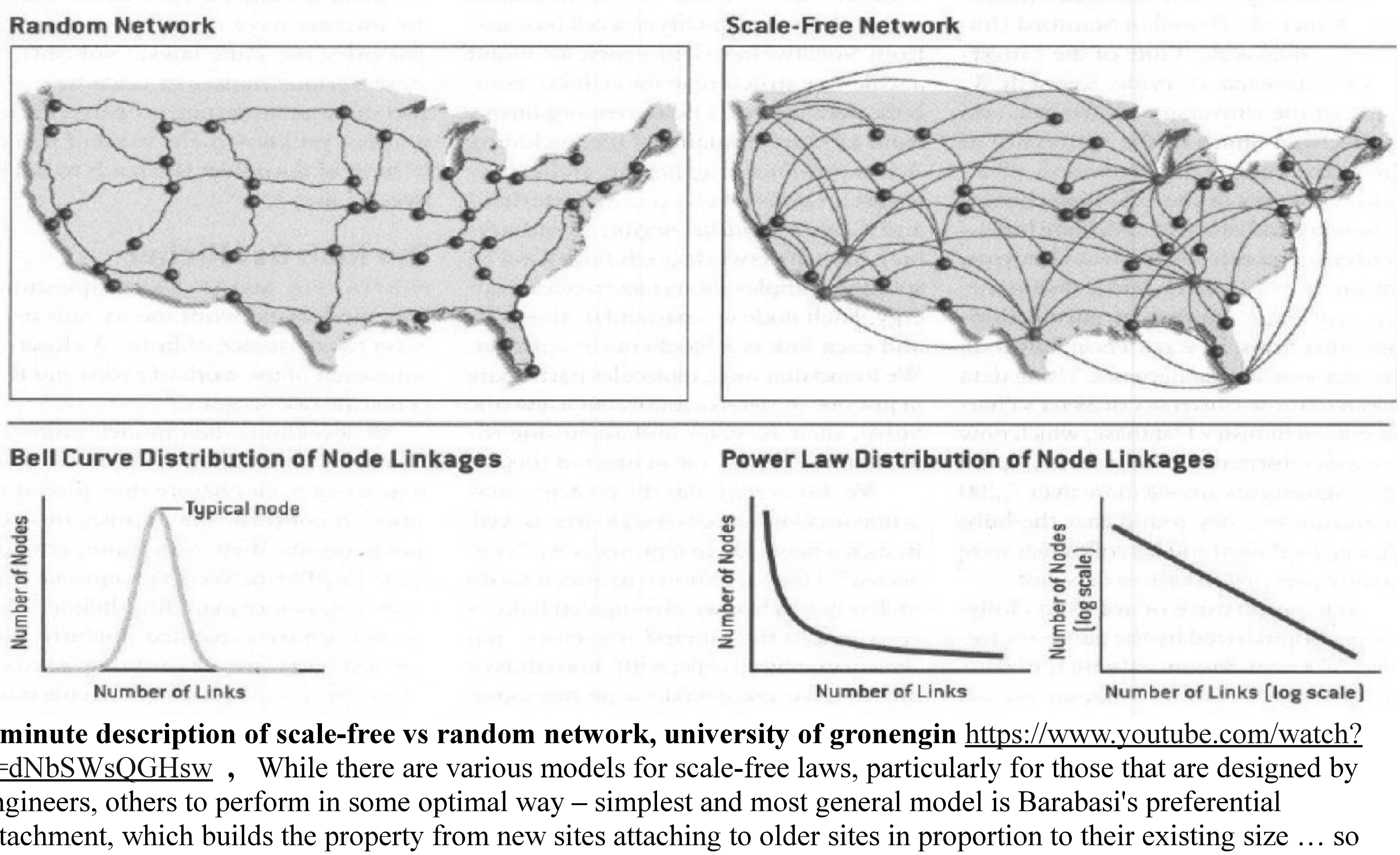
SMALL worlds are not everywhere. Brain scientists find there is more than small world in brain cells. Bassett and Bulmore (2017) **Small-World Brain Networks Revisited (**The Neuroscientist 2017, Vol. 23(5) 499–516) : "... small-worldness—the combination of non-random clustering with near-random path length— has been very frequently reported across a wide range of neuro-science studies. Small-world topology has been highly replicated across multiple species and scales from structural and functional MRI studies of large-scale brain networks in humans to multielectrode array recordings of cellular networks in cultures (Bettencourt and others 2007) and intact animals (van den Heuvel and others 2016). About 3 to 4 years ago, an important series of papers began to be published that could be regarded as "black swans" refuting the general importance of small-worldness in an understanding of brain networks. … Data that reveal high-density cortical graphs in which economy of connections is achieved by weight heterogeneity and distance-weight correlations. Recent connectomic tract tracing reveals that, contrary to previous thought, the cortical inter-areal network has high density. This finding leads to a necessary revision of the relevance of some of the graph theoretical notions.

MAIN AREA II BEYOND RANDOM MODEL: POWER LAWS IN NETWORKS – the SCALE-FREE DEBATE?

Networks can generate power laws for the number of nodes with different linkages. Barabasi et al posit a growth process with new vertices added each period that have a higher probability to connect to smaller vertexes. Probability of connecting to NY website is proportional to # connections to NY/ All links.



(a) Random network (b) Scale-free network



4 minute description of scale-free vs random network, university of gronengin https://www.youtube.com/watch?v=dNbSWsQGHsw , While there are various models for scale-free laws, particularly for those that are designed by engineers, others to perform in some optimal way – simplest and most general model is Barabasi's preferential attachment, which builds the property from new sites attaching to older sites in proportion to their existing size … so big get bigger over time. But the power laws – scale free features – are imperfect descriptors of reality.

Great value of scale free is you learn about properties/attributes of rare "big events" from properties of small events. But even modest deviation at upper tail says cannot trust patterns based on big data for smaller events to replicate for key events.

Since lots of small world networks in empirical data, to what extent, if at all, does node linkages of a small world graph fit power law distribution? Small world network does not imply a power law just as competitive economy does not mean levels of inequality we observe are "necessary" for growing economy. In basic form the small world model has a degree distribution that is roughly normal, very different from observed skewed distributions; whereas the preferential attachment model has low path lengths and a heavy-tailed degree distribution,but ow clustering, and degree distribution that fits observed data only up to a point. So one answer is no! (http://allendowney.blogspot.com/2016/09/its-small-world-scale-free-network.html)

The brainstem reticular formation is a small-world, BUT not scale-free, network (Humphries, Gurney and. Prescott in Proc. R. Soc. B (2006) 273, 503–511: … several complex systems may have simple graph-theoretic characterizations as so-called 'small-world' and 'scale-free' networks. We conclude that the medial RF is configured to create small-world (implying coherent rapid-processing capabilities), but not scale-free, type networks under assumptions which are amenable to quantitative measurement.

But some networks appear to fit both small world and power law: **Complex Networks: Small-World,Scale-Free and Beyond** Xiao Fan Wang and Guanrong Chen http://rakaposhi.eas.asu.edu/cse494/ scalefree.pdf

Table 1.
Small-world pattern and scale-free property of several real networks. Each network has the number of nodes N, the clustering coefficient C, the average path length L and the degree exponent γ of the power-law degree distribution. The WWW and metabolic network are described by directed graphs.

Network	Size	Clustering coefficient	Average path length	Degree exponent
Internet, domain level [13]	32711	0.24	3.56	2.1
Internet, router level [13]	228298	0.03	9.51	2.1
WWW [14]	153127	0.11	3.1	$\gamma_{in} = 2.1 \quad \gamma_{out} = 2.45$
E-mail [15]	56969	0.03	4.95	1.81
Software [16]	1376	0.06	6.39	2.5
Electronic circuits [17]	329	0.34	3.17	2.5
Language [18]	460902	0.437	2.67	2.7
Movie actors [5, 7]	225226	0.79	3.65	2.3
Math. co-authorship [19]	70975	0.59	9.50	2.5
Food web [20, 21]	154	0.15	3.40	1.13
Metabolic system [22]	778	–	3.2	$\gamma_{in} = \gamma_{out} = 2.2$

Downey offers a model that links power law and small world https://github.com/AllenDowney/ThinkComplexity2/blob/master/code/fof_model.ipynb

To get some fit must modify models with an extra parameter. If parameter differs markedly across comparable phenomenon lost a lot of the virtue of simple relation.

Is Scale-free property as widespread as claimed?

Big and not quite unpleasant debate about this.

NO IT IS NOT A central claim in network science is that real-world networks are typically scale free,meaning that the fraction of nodes with degree k follows a power law, decaying like k^{-a}, often with 2 < < 3. But **Scale-free networks are rare. Broido, A.D., Clauset, A. Nat Commun 10, 1017 (2019). :** We test the universality of scale-free structure by applying state-of-the-art statistical tools to a large corpus of nearly 1000 network data sets drawn from social, biological,technological, and informational sources. We fit the power-law model to each degree distribution,test its statistical plausibility, and compare it via a likelihood ratio test to alternative, non-scale-free models, e.g., the log-normal. Across domains, we find that scale-free networks are rare, with only 4% exhibiting the strongest-possible evidence of scale-free structure and 52% exhibiting the weakest-possible evidence. Furthermore, evidence of scale-free structure is not uniformly distributed across sources: social networks are at best weakly scale free, while a handful of technological and biological networks can be called strongly scale free. These results undermine the universality of scale-free networks and reveal that real-world networks exhibit a rich structural diversity that will likely require new ideas and mechanisms to explain.

Table 1 Comparison of scale-free and alternative distributions

		Test outcome		
Alternative	$p(x) \propto f(x)$	M_{PL}	**Inconclusive**	M_{Alt}
Exponential	$e^{-\lambda x}$	33%	26%	41%
Log-normal	$\frac{1}{x} e^{-\frac{(\log x-\mu)^2}{2\sigma^2}}$	12%	40%	48%
Weibull	$e^{-(\frac{x}{b})^a}$	33%	20%	47%
Power law with cutoff	$x^{-\alpha} e^{-\lambda x}$	–	44%	56%

The percentage of network data sets that favor the power-law model M_{PL}, alternative model M_{Alt}, or neither, under a likelihood-ratio test, along with the form of the alternative distribution $f(x)$

YES IT IS AS WIDESPREAD AS CLAIMED: Barabasi "https://www.barabasilab.com/post/love-is-all-you-need "before accepting the attention-grabbing claims of the BC paper, you must peek under its hood. And if you take the time to do that, you will find ... a fictional criterion of scale-free networks. Most important, you will find that their central criterion fails the most elementary tests."
"they fail to see the scale-free property because they invent a new criterion of weak and strong scale-free networks. And the real surprise? Even the exact model of scale-free networks, following a pure power law, fails their test. As incredible as this sounds, all the evidence is in Appendix E, on the very last page. So let's dig into it."

Note: no disagreement that networks are based on distributions with heavy tails, with power-like characteristics at the upper end of distributions. Argument is about distributions that look non-power law for a while and then get heavy tail, which often is heavier than power law predicts: Is this power law like or not?

YES IT IS BASED ON "WELL-DONE" STATISTICAL ANALYSIS based on extreme value theory – statistical model of the extremes of distributions.

What is extreme value theory? Statistics model of extremes used in risk analysis, financial economics, to deal with events in far tail. It turns out that there are only three families of EVDs. They are parameterized by a real number ξ, called the extreme value index. The three families are Weibull with $\xi < 0$, Gumbel with $\xi = 0$, and Fréchet with $\xi > 0$.

IVoitalov, P van der Hoorn, R, van der Hofstad and D Krioukov **Scale-free networks well done PHYSICAL REVIEW RESEARCH 1, 033034 (2019)**

We bring rigor to detecting power laws in empirical degree distributions in real-world networks.

We define power-law distributions as part of regularly varying distributions … used in statistics and other fields that allows the **distribution to deviate from a pure power law but without affecting the power-law tail exponent.** .. and identify three estimators of these exponents that are statistically consistent—that is, converging to the true value of the exponent for any regularly varying distribution

In a representative collection of synthetic and real-world data. According to their estimates, real-world scale-free networks are not as rare as indicated by the popular unrealistic assumption that real-world data come from power laws of pristine purity, void of noise, and deviations. Allowing for real-world impurity via **regularly varying distributions,** we find that 49% of the considered **undirected** networks have degree sequences that are power-law. That 24% of **directed network**s have both in- and out degree sequences that are power-law, while 82% have either in- or out-degree sequence that is power-law. and that 35% of **bipartite** networks have power-law degree sequences for both types of nodes, while in 74% have least one type of nodes has a power-law degree sequence.

Key to analysis is definition of when a distribution is **approximately** a power law. If it is regularly varying with $F(k) = l(k)\ k^{-\alpha}$ where $\alpha > 0$, and $l(k)$ is a slowly varying function. This formalize the point that the distribution exhibits a power-law tail at high degrees, but has an arbitrary shape at small degrees.

Key to empirical analysis of the tail exponent of this power law from the measured degree sequence requires that as the number of samples increases, the estimated values of the exponent to **converge to the true exponent value** ξ regardless of the slowing-varying function (k). do this by estimating the extreme value index ξ of the distribution. There are only three ways to estimate the extreme value index – The Hill moments, and kernel estimators – for which the double bootstrap method that automatically determines the optimal value of κ is proven to be both optimal and consistent.

TABLE I. The tail exponent estimation results for the 35 real-world undirected networks collected from the KONECT database [56]. Each network name is followed by its KONECT code in braces. The estimators return estimates $\hat{\xi}$ of ξ that are translated to $\hat{\gamma} = 1 + 1/\hat{\xi}$. If $\hat{\xi} \leqslant 0$, then $\hat{\gamma}$ is set to ∞. The estimates are colored according to the definitions in Sec. V: (1) not power-law networks, at least one estimate is nonpositive $\hat{\xi} \leqslant 0$ (red); (2) hardly power-law networks, all the estimates are positive $\hat{\xi} > 0$ and at least one estimate is $\hat{\xi} \leqslant 1/4$ (yellow); (3) power-law networks with a divergent second moment, all the estimates are $\hat{\xi} > 1/2$ (green); and (4) other power-law networks, the rest of the cases (blue).

Network name	n	$\bar{k}$	$\hat{\gamma}^{\text{Hill}}$	$\hat{\gamma}^{\text{mom}}$	$\hat{\gamma}^{\text{kern}}$
CAIDA (IN)	26 475	4.03	2.1	2.11	2.11
Skitter (SK)	1 696 415	13.08	2.38	2.36	2.43
Actor collaborations (CL)	382 219	173.28	3.71	6.7×10^3	2.36
Amazon (CA)	334 863	5.53	3.99	3.48	3.44
arXiv (AP)	18 771	21.1	4.41	5.78	7.29
Bible names (MN)	1773	10.3	3.09	3.36	2.88
Brightkite (BK)	58 228	7.35	3.51	3.8	2.96
Catster (Sc)	149 684	72.8	2.09	2.06	1.98
Catster/Dogster (Scd)	623 748	50.33	2.1	2.11	2.04
Chicago roads (CR)	1467	1.77	77.92	∞	∞
DBLP (CD)	317 080	6.62	6.59	13.99	3.06
Dogster (Sd)	426 816	40.03	2.15	2.15	2.12
Douban (DB)	154 908	4.22	4.42	6.88	1.86
U. Rovira I Virgili (A@)	1133	9.62	6.49	∞	∞
Euro roads (ET)	1174	2.41	4.73	44.48	29.57
Flickr (LF)	1 715 254	18.13	3.94	4.29	5.02
Flickr (FI)	105 938	43.74	6.18	1.79	1.65
Flixster (FX)	2 523 386	6.28	53.63	1.93	1.95
Gowalla (GW)	196 591	9.67	2.8	2.8	2.86
Hamsterster (Shf)	1858	13.49	4.45	8.09	3.51
Hamsterster (Sh)	2426	13.71	4.57	25.39	6.32
Hyves (HY)	1 402 673	3.96	2.98	2.23	1.99
LiveJournal (Lj)	5 203 763	18.72	3.86	4.04	3.15
Livemocha (LM)	104 103	42.13	9.13	∞	2.39
Orkut (OR)	3 072 441	76.28	3.58	2.65	3.35
Power grid (UG)	4941	2.67	6.62	7.76	9.2

Hmm. Holme, P. Rare and everywhere: Perspectives on scale-free networks. *Nat Commun* **10,** 1016 (2019)

Now we have one camp of network scientists thinking of scale free networks as ideal objects in the large-size limit, and another seeing them as concrete objects belonging to the real world. Scientists often sneer at the humanities with their schools of thought and lack of consensus, but such is the current state of network science.....

The most common arguments against the claim that scale-free networks are rare relate to the concept's origins in complexity science and in particular the fact that scale-freeness is only well-defined in the infinite-size limit....a network is scale-free if its degree distribution approaches a power law as the network keeps growing following the same mechanisms. Thus, fluctuations that make a network fail a statistical test would not matter, because if one let the network Evolve, it could soon pass the test...

A celebrated result states that scale-free networks with degree exponents less than three don't have epidemic thresholds --epidemics can always reach a large fraction of the population – (ie need very fast fall-off of tails or spreads widely). But this holds only for infinite networks. If, BC estimate that a finite network is consistent with a powerlaw degree distribution of exponent two, then it does not have an epidemic threshold.

(The main point) is that "knowledge of whether or not a distribution is heavy-tailed is far more important than whether it can be fit using a power law"-- Stumpf, M. P. H. & Porter, M. A. Critical truths about power laws. Science 335, 665–666 (2012) Still, it often feels like the topic of scale-free networks transcends science !!!

I like the extreme value perspective.