Complex Network Theory

Lecture 7

Scale free networks

Instructor: S. Mehdi Vahidipour
(Vahidipour@kashanu.ac.ir)
Spring 2017

Thanks A. Rezvanian
A. Barabasi, L.Adamic,

Outline

- Heavy Tail distributions
- Power law distributions
- Scale free networks
- 20/80 rule
- What kinds of processes generate power laws?

- Next class:
  - Community structure

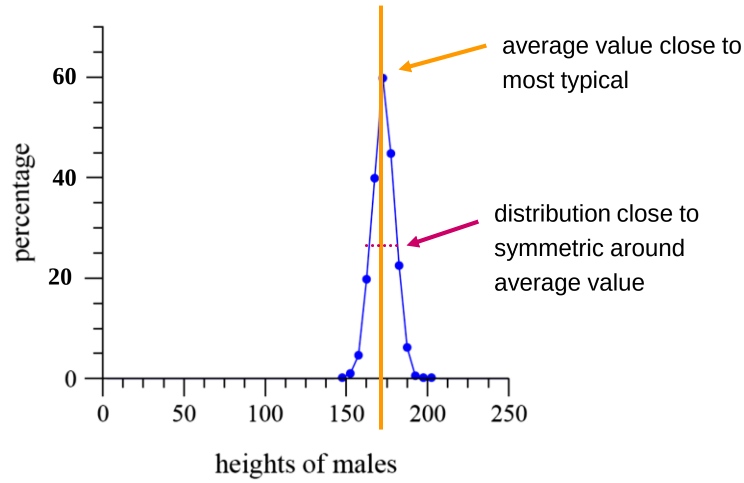
What is a heavy tailed-distribution?

- Right skew
  - Normal distribution (not heavy tailed)
    - e.g. heights of human males: centered around 180cm
  - Zipf's or power-law distribution (heavy tailed)
    - e.g. city population sizes: Tehran 12 million, but many, many small towns
- High ratio of max to min
  - human heights
    - tallest man: 272cm, shortest man: 56 cm *ratio: 4.8*
      from the Guinness Book of world records
  - city sizes
    - Tehran: pop. 12 million, a village pop. 78, *ratio: 150,000*

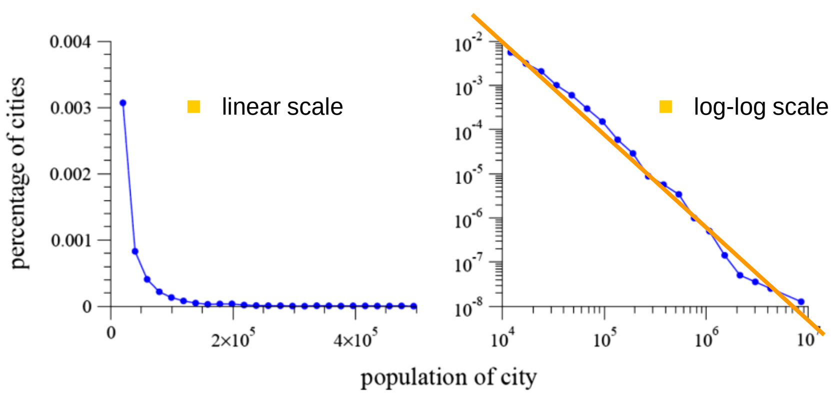
The Heavy Tail

- The power law distribution implies an "infinite variance"
  - The "area" of "big ks" in an exponential distribution tend to zero with $k \to \infty$
  - This is not true for the power law distribution, implying an infinite variance
- In other words, the power law implies that
  - The probability to have elements very far from the average is not negligible
- Using an exponential distribution
  - The probability for a Web page to have more than 100 incoming links, considering the average number of links for page, would be less in the order of 1^{-20}
  - which contradicts the fact that we know a lot of "well linked" sites

Normal (also called Gaussian) distribution of human heights

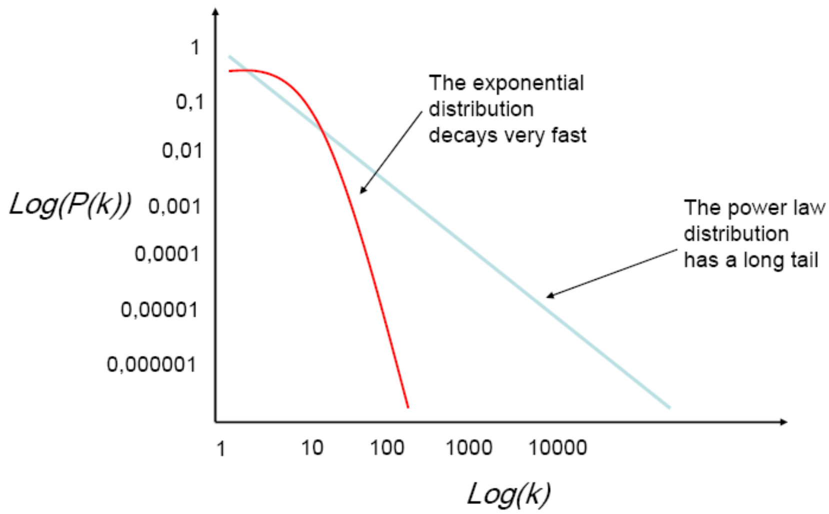


Power-law distribution



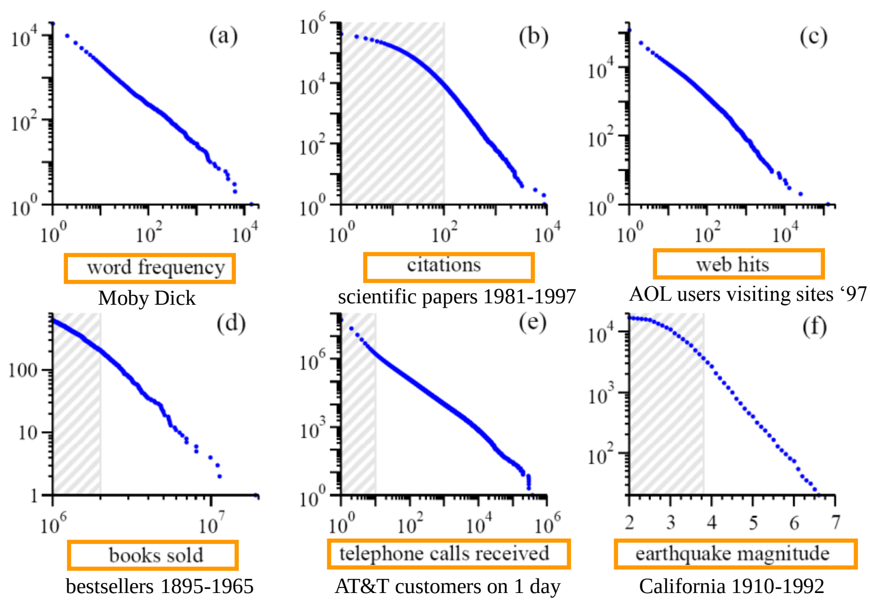
- high skew (asymmetry)
- straight line on a log-log plot

Power-law vs. Exponential distribution



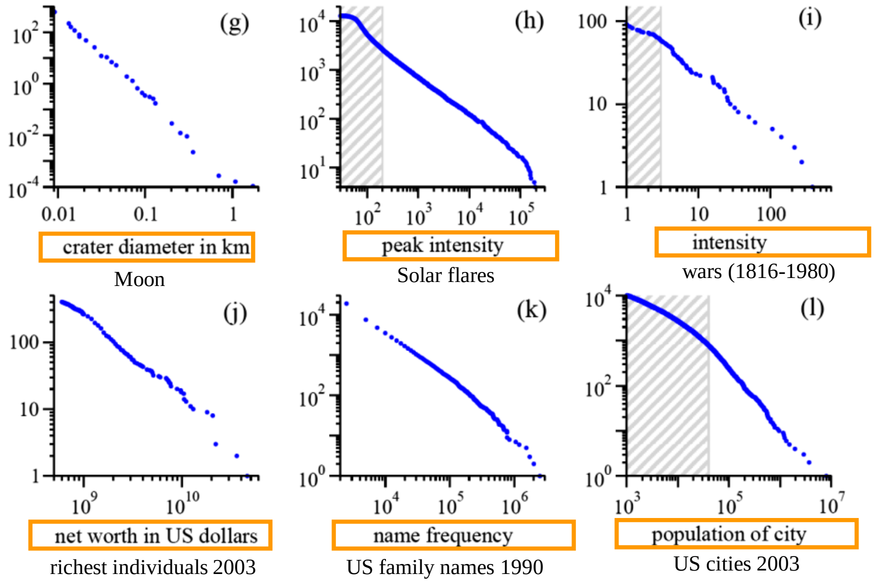
Power laws are seemingly everywhere

note: these are cumulative distributions, more about this in a bit…



Source: MEJ Newman, 'Power laws, Pareto distributions and Zipf's law', Contemporary Physics 46, 323–351 (2005)

Yet more power laws



Source: MEJ Newman, 'Power laws, Pareto distributions and Zipf's law', Contemporary Physics 46, 323–351 (2005)

The Power-law in real networks

Average k — Power law exponents

Network	Size	$\langle k \rangle$	κ	γ_{out}	γ_{in}	ℓ_{real}	ℓ_{rand}	ℓ_{pow}	Reference
WWW	325 729	4.51	900	2.45	2.1	11.2	8.32	4.77	Albert, Jeong, and Barabási 1999
WWW	4×10^7	7		2.38	2.1				Kumar *et al.*, 1999
WWW	2×10^8	7.5	4000	2.72	2.1	16	8.85	7.61	Broder *et al.*, 2000
WWW, site	260 000				1.94				Huberman and Adamic, 2000
Internet, domain*	3015–4389	3.42–3.76	30–40	2.1–2.2	2.1–2.	4	6.3	5.2	Faloutsos, 1999
Internet, router*	3888	2.57	30	2.48	2.48	12.15	8.75	7.67	Faloutsos, 1999
Internet, router*	150 000	2.66	60	2.4	2.4	11	12.8	7.47	Govindan, 2000
Movie actors*	212 250	28.78	900	2.3	2.3	4.54	3.65	4.01	Barabási and Albert, 1999
Co-authors, SPIRES*	56 627	173	1100	1.2	1.2	4	2.12	1.95	Newman, 2001b
Co-authors, neuro.*	209 293	11.54	400	2.1	2.1	6	5.01	3.86	Barabási *et al.*, 2001
Co-authors, math.*	70 975	3.9	120	2.5	2.5	9.5	8.2	6.53	Barabási *et al.*, 2001
Sexual contacts*	2810			3.4	3.4				Liljeros *et al.*, 2001
Metabolic, *E. coli*	778	7.4	110	2.2	2.2	3.2	3.32	2.89	Jeong *et al.*, 2000
Protein, *S. cerev.**	1870	2.39		2.4	2.4				Jeong, Mason, *et al.*, 2001
Ythan estuary*	134	8.7	35	1.05	1.05	2.43	2.26	1.71	Montoya and Solé, 2000
Silwood Park*	154	4.75	27	1.13	1.13	3.4	3.23	2	Montoya and Solé, 2000
Citation	783 339	8.57			3				Redner, 1998
Phone call	53×10^6	3.16		2.1	2.1				Aiello *et al.*, 2000
Words, co-occurrence*	460 902	70.13		2.7	2.7				Ferrer i Cancho and Solé, 2001
Words, synonyms*	22 311	13.48		2.8	2.8				Yook *et al.*, 2001b

The Ubiquity of the Power Law

- The previous table includes not only technological networks
  - Most real systems and events have a probability distribution that
    - Does not follow the "normal" distribution
    - and obeys to a power law distribution
- Examples, in addition to technological and social networks
  - The distribution of size of files in file systems
  - The distribution of network latency in the Internet
  - The networks of protein interactions (a few protein exists that interact with a large number of other proteins)
  - The power of earthquakes: statistical data tell us that the power of earthquakes follow a power-law distribution
  - The size of rivers: the size of rivers in the world is power law
  - The size of industries, i.e., their overall income
  - The richness of people
  - In these examples, the exponent of the power law distribution is always around 2.5
- The power law distribution is the "normal" distribution for complex systems (i.e., systems of interacting autonomous components)
  - We see later how it can be derived…

The 20-80 Rule

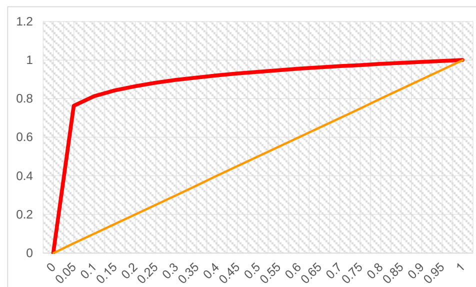
- It's a common "way of saying"
  - But it has scientific foundations
  - For all those systems that follow a power law distribution
- Examples
  - The 20% of the Web sites gests the 80% of the visits (actual data: 15%-85%)
  - The 20% of the Internet routers handles the 80% of the total Internet traffic
  - The 20% of world industries hold the 80% of the world's income
  - The 20% of the world population consumes the 80% of the world's resources
  - The 20% of the Italian population holds the 80% of the lands (that was true before the Mussolini fascist regime, when lands redistribution occurred)
  - The 20% of the earthquakes caused the 80% of the victims
  - The 20% of the rivers in the world carry the 80% of the total sweet water
  - The of the proteins handles the of the most critical metabolic processes
- Does this derive from the power law distribution? YES!

80/20 rule

- The fraction W of the wealth in the hands of the richest P of the population is given by

$$W = P^{(\alpha-2)/(\alpha-1)}$$

- Example: US wealth: $\alpha = 2.1$
  - richest 20% of the population holds 86% of the wealth

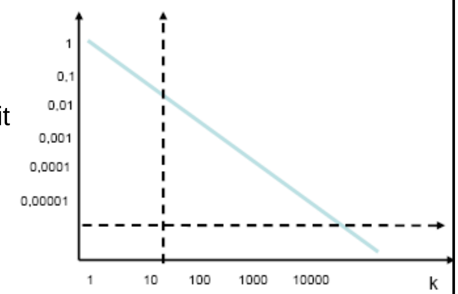


Hubs and Connectors

- Scale free networks exhibit the presence of nodes that
  - Act as hubs, i.e., as point to which most of the other nodes connects to
  - Act as connectors, i.e., nodes that make a great contributions in getting great portion of the network together
  - “smaller nodes” exists that act as hubs or connectors for local portion of the network
- This may have notable implications, as detailed below

Why "Scale-Free" Networks

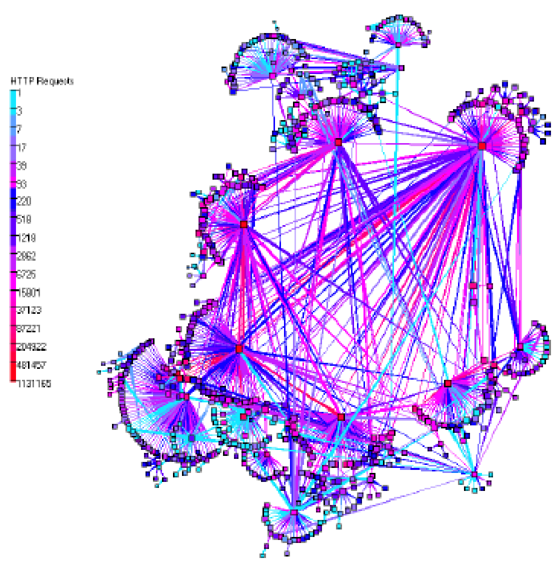
- Why networks following a power law distribution for links are called "**scale free**"?
  - Whatever the scale at which we observe the network
  - The network looks the same, i.e., it looks similar to itself
- The overall properties of the network are preserved independently of the scale



- In particular
  - If we cut off the details of a network – skipping all nodes with a number of links the limited – network will preserve its power-law structure
  - If we consider a sub-portion of any network it have the network, will same overall structure of the whole network

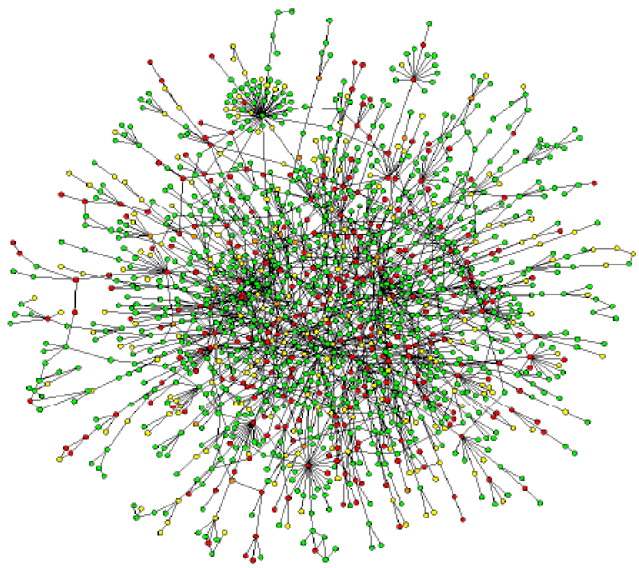
How do they look like?

Web Cache Network



How do they look like?

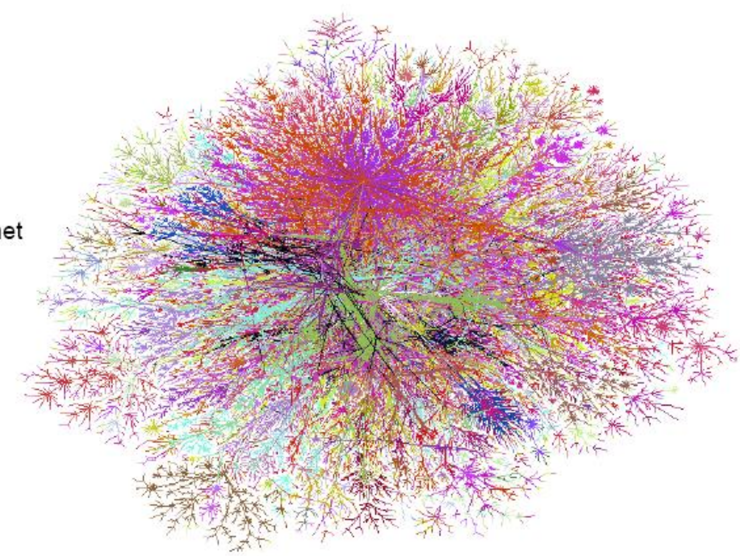
Protein Network



Complex Network Theory, S. M. Vahidipour, Spring 2017.

How do they look like?

The Internet Routers



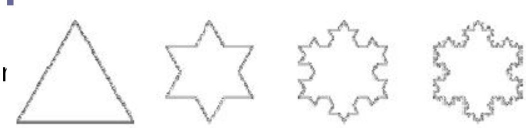
Complex Network Theory, S. M. Vahidipour, Spring 2017.

Fractals and Scale Free Networks

- The nature is made up of mostly "fractal objects"
  - The fractal term derives from the fact that they have a non-integer dimension
  - 2-d objects have a "size" (i.e., a surface) that scales with the square of the linear size A=kL2
  - 3-d objects have a "size" (i.e., a volume) that scales with the cube of the linear size V=kL3
  - Fractal objects have a "size" that scales with some fractions of the linear size S=kLa/b
- Fractal objects have the property of being "self-similar" or "scale-free"
  - Their "appearance" is independent from the scale of observation
  - They are similar to itself independently of whether you look at the from near and from far
  - That is, they are scale-free

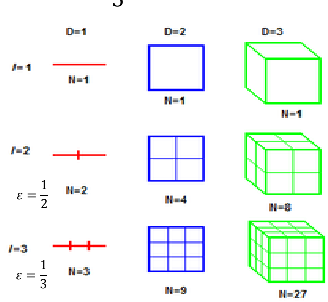
Examples of Fractals

- The Koch snowflake
  - Coastal Regions & River systems
  - Lymphatic systems
  - The distribution of masse in the universe



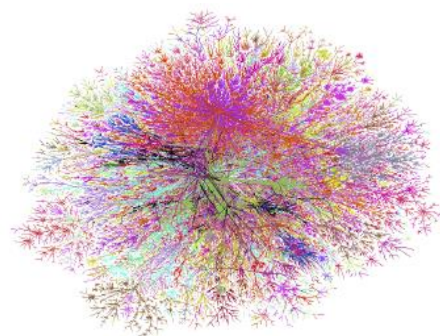
$$\log_{\varepsilon} N = -D = \frac{\log N}{\log \varepsilon}$$

$$\varepsilon = \frac{1}{3}, N = 4 \Rightarrow D = 1.2619$$



Scale Free Networks are Fractals?

- Yes, in fact:
  - They are the same at whatever dimension we observe them
  - Also, the fact that they grow according to a power law can be considered as a sort of fractal dimension of the network…
- Having a look at the figures clarifies the analogy



Power law distribution

- Straight line on a log-log plot

$$\ln(p(x)) = c - \alpha \ln(x)$$

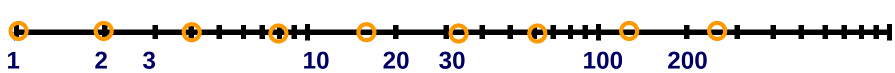
- Exponentiate both sides to get that *p(x)*, the probability of observing an item of size 'x' is given by

$$p(x) = Cx^{-\alpha}$$

power law exponent α

Normalization constant (probabilities over all *x* must sum to 1)

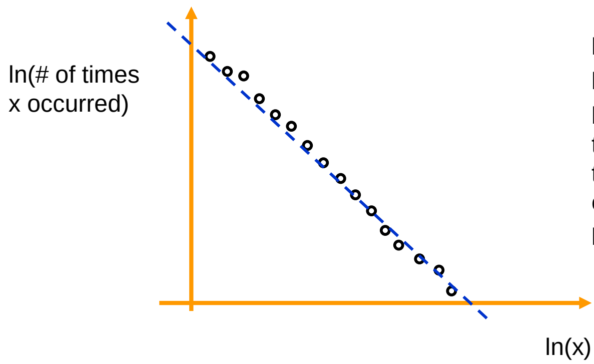
- powers of a number will be uniformly spaced (Logarithmic axes)



- $2^0=1$, $2^1=2$, $2^2=4$, $2^3=8$, $2^4=16$, $2^5=32$, $2^6=64$,….

Fitting power-law distributions

- Most common and not very accurate method:
  - Bin the different values of x and create a frequency histogram



ln(x) is the natural logarithm of x, but any other base of the logarithm will give the same exponent of a because $\log_{10}(x) = \ln(x)/\ln(10)$

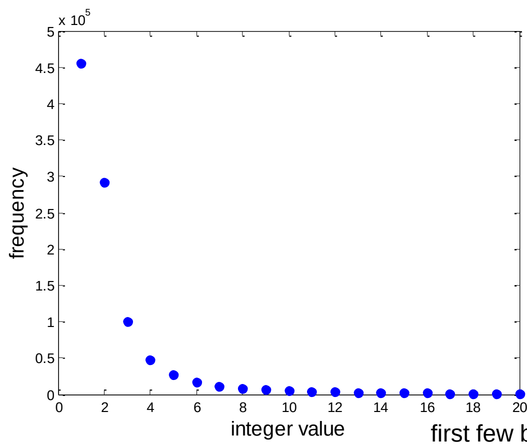
x can represent various quantities, the indegree of a node, the magnitude of an earthquake, the frequency of a word in text

Example on an artificially generated data set

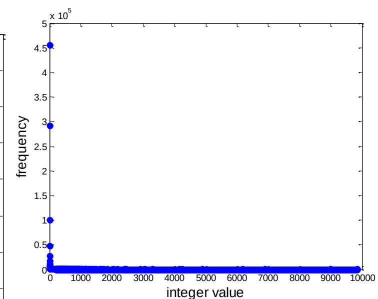
- Take 1 million random numbers from a distribution with $\alpha = 2.5$
- Can be generated using the so-called 'transformation method'
  - Generate random numbers r on the unit interval $0 \leq r < 1$
  - then $x = (1-r)^{-1/(\alpha-1)}$ is a random power law distributed real number in the range $1 \leq x < \infty$

Linear scale plot of straight bin of the data

- How many times did the number 1 or 3843 or 99723 occur
- Power-law relationship not as apparent
- Only makes sense to look at smallest bins



first few bins

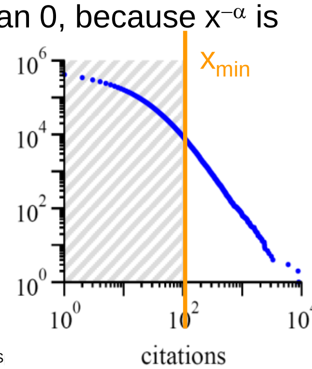


whole range

Where to start fitting?

- some data exhibit a power law only in the tail
- after binning or taking the cumulative distribution you can fit to the tail
- so need to select an x_{min} the value of x where you think the power-law starts
- certainly x_{min} needs to be greater than 0, because $x^{-\alpha}$ is infinite at x = 0

Example: Distribution of citations to papers where power law is evident only in the tail (x_{min} > 100 citations)



Source: MEJ Newman, 'Power laws, Pareto distributions and Zipf's law', Contemporary Physics **46**, 323–351 (2005)

Some exponents for real world data

	x_{min}	exponent α
frequency of use of words	1	2.20
number of citations to papers	100	3.04
number of hits on web sites	1	2.40
copies of books sold in the US	2 000 000	3.51
telephone calls received	10	2.22
magnitude of earthquakes	3.8	3.04
diameter of moon craters	0.01	3.14
intensity of solar flares	200	1.83
intensity of wars	3	1.80
net worth of Americans	$600m	2.09
frequency of family names	10 000	1.94
population of US cities	40 000	2.30

Many real world networks are power law

	exponent α (in/out degree)
film actors	2.3
telephone call graph	2.1
email networks	1.5/2.0
sexual contacts	3.2
WWW	2.3/2.7
internet	2.5
peer-to-peer	2.1
metabolic network	2.2
protein interactions	2.4

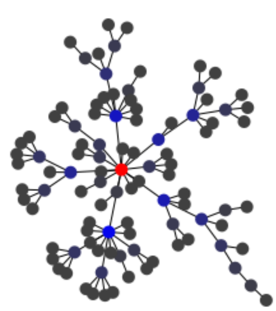
Complex Network Theory, S. M. Vahidipour, Spring 2017.

Preferential Attachment in Networks

- First considered by [Price 65] as a model for citation networks
  - each new paper is generated with m citations (mean)
  - new papers cite previous papers with probability proportional to their indegree (citations)
  - what about papers without any citations?
    - each paper is considered to have a “default” citation
- Power law with exponent

generating power-law networks

- Nodes appear over time (growth)
- Nodes prefer to attach to nodes with many connections (preferential attachment, cumulative advantage)

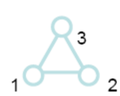


Barabási-Albert model (BA model)

- Undirected model: each node connects to other nodes with probability proportional to their degree
  - the process starts with some initial subgraph (m_0 all-all connected node)
  - each node comes with m edges
  - the probability of tipping the new nodes to the old ones is proportional to the degrees of old nodes is a kind of preferential attachment algorithm
  - After t time steps, the network will have $n=t+m_0$ nodes and $M=m_0+mt$ edges
- It can be shown that this leads to a power law network!

Basic BA-model

- Very simple algorithm to implement
  - start with an initial set of m_0 fully connected nodes
    - e.g. $m_0 = 3$

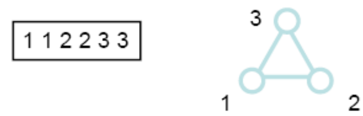


1 1 2 2 2 3 3 4 5 6 6 7 8 ....

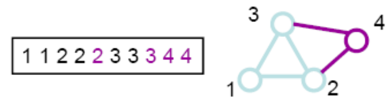
  - now add new vertices one by one, each one with exactly m edges
  - each new edge connects to an existing vertex in proportion to the number of edges that vertex already has → **preferential attachment**
  - easiest if you keep track of edge endpoints in one large array and select an element from this array at random
    - the probability of selecting any one vertex will be proportional to the number of times it appears in the array – which corresponds to its degree

Generating BA graphs

- To start, each vertex has an equal number of edges (2)
  - the probability of choosing any vertex is 1/3

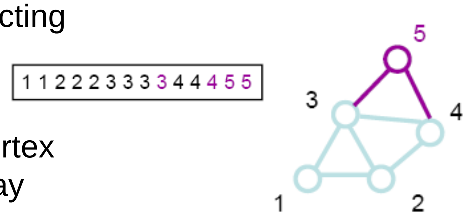


- We add a new vertex, and it will have m edges, here take m=2
  - draw 2 random elements from the array – suppose they are 2 and 3



- Now the probabilities of selecting 1,2,3,or 4 are

  1/5, 3/10, 3/10, 1/5



- Add a new vertex, draw a vertex for it to connect from the array
- etc.

Proof of the scale-freeness

- Assume for simplicity that the degree k_i for any node i is a continuous variable
- The probability of the tipping a node to node i is

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

- Because of the assumptions, k_i is expected to grow proportionally to $\Pi(k_i)$, that is to its probability of having a new edge
- Consequently, and because m edges are attached at each time, k_i should obey the differential equation aside

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_{j=1}^{n-1} k_j}$$

Proof of the scale-freeness

- The sum $\sum_{j=1}^{n-1} k_j$
- Goes over all nodes except the new ones
- This it results in

$$\sum_{j=1}^{n-1} k_j = 2mt - m$$

- Remember that the total number of edges is almost mt and that here is edge is twice
- Substituting in the differential equation

$$\frac{\partial k_i}{\partial t} = m\frac{k_i}{\sum_{j=1}^{n-1} k_j} = m\frac{k_i}{2mt - m} \approx \frac{k_i}{2t}$$

Proof of the scale-freeness

- We have now to solve this equation $\frac{\partial k_i}{\partial t} = \frac{k_i}{2t}$
  - That is, we have find a k_i(t) function such as its derivative is equal to itself, divided by 2t
- We now show this is:

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}; \quad \text{with } \beta = \frac{1}{2}$$

- In fact:

$$\frac{\partial}{\partial t}\left(m\left(\frac{t}{t_i}\right)^{\beta}\right) = \frac{1}{2}\frac{m}{t_i^{\beta}}\frac{1}{t^{\beta}} = \frac{1}{2}\frac{m}{t_i^{\beta}}\frac{1}{t^{\beta}}\frac{t^{\beta}}{t^{\beta}} = \frac{m}{2}\frac{t^{\beta}}{t_i^{\beta}}\frac{1}{t^{2\beta}} = \frac{k_i(t)}{2t}$$

- where we also consider the initial condition $k_i(t_i)$=m, where ti is the time at which node i has arrived

Proof of the scale-freeness

- The $k_i(t)$ function that we have not calculated shows that the degree of each node grown with a power law with time
- Now, let's calculate the probability that a node has a degree $k_i(t)$ smaller than k
- We have

$$P[k_i(t) < k] = P\left[m\frac{t^\beta}{t_i^\beta} < k\right] = P\left[m^{\frac{1}{\beta}}\frac{t^{\beta\frac{1}{\beta}}}{t_i^{\beta\frac{1}{\beta}}} < k^{\frac{1}{\beta}}\right] =$$

$$= P\left[m^{\frac{1}{\beta}}\frac{t}{t_i} < k^{\frac{1}{\beta}}\right] = P\left[t_i > \frac{m^{\frac{1}{\beta}}t}{k^{\frac{1}{\beta}}}\right]$$

Proof of the scale-freeness

- Now let's remember that we add nodes at each time interval
- Therefore, the probability t_j for a node, that is the probability for a node to have arrived at time t_i is a constant and is:

$$P(t_i) = \frac{1}{t + m_0}$$

- Substituting this into the previous probability distribution

$$P[k_i(t) < k] = P\left[t_i > \frac{m^{\frac{1}{\beta}}t}{k^{\frac{1}{\beta}}}\right] = 1 - P\left[t_i \le \frac{m^{\frac{1}{\beta}}t}{k^{\frac{1}{\beta}}}\right] = 1 - \frac{m^{\frac{1}{\beta}}t}{k^{\frac{1}{\beta}}(t + m_0)}$$

Proof of the scale-freeness

- Now given the probability distribution

$$P[k_i(t) < k]$$

- Which represents the probability that a node i has less than k link

$$P(k) = \frac{\partial P[k_i(t) < k]}{\partial k}$$

- The probability that a node has exactly k link can be derived by the derivative of the probability distribution

$$P(k) = \frac{\partial P[k_i(t) < k]}{\partial k} = \frac{\partial}{\partial k}\left(1 - \frac{m^{\frac{1}{\beta}} t}{k^{\frac{1}{\beta}}(t + m_0)}\right) = \frac{2m^{\frac{1}{\beta}} t}{m_0 + t} \frac{1}{k^{\frac{1}{\beta}+1}}$$

Conclusion of the Proof

- Given P(k):

$$P(k) = \frac{2m^{\frac{1}{\beta}} t}{m_0 + t} \frac{1}{k^{\frac{1}{\beta}+1}}$$

- After a while, that is for t→∞

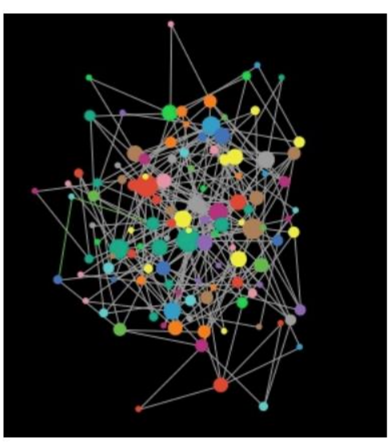
$$P(k) \approx 2m^{\frac{1}{\beta}} k^{-\frac{1}{\beta}-1} = 2m^{\frac{1}{\beta}} k^{-\gamma} \quad \text{where } \gamma = \frac{1}{\beta} + 1 = 3$$

- **we have obtained a power law probability density,** with an exponent which is independent of any parameter (being the only initial parameter m)

Generality of the BA Model

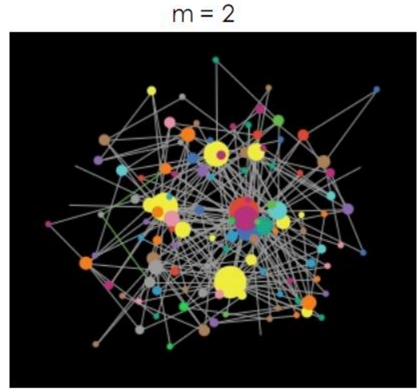
- In its simplicity, the BA model captures the essential characteristics of a number of phenomena
  - In which events determining "size" of the individuals in a network
  - Are not independent from each other
  - Leading to a power law distribution
- So, it can somewhat explain why the power law distribution is as ubiquitous as the normal Gaussian distribution
- Examples
  - **Gnutella (the first decentralized P2P network)**: a peer which has been there for a long time, has already collected a strong list of acquaintances, so that any new node has higher probability of getting aware of it
  - **Rivers**: the eldest and biggest a river, the more it has probability to break the path of a new river and get its water, thus becoming even bigger
  - **Industries**: the biggest an industry, the more its capability to attract clients and thus become even bigger
- **Richness**: the rich I am, the more I can exploit my money to make new money $\rightarrow$ "RICH GET RICHER"

random non-preferential and preferential growth



random

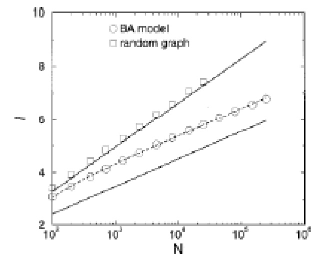
m = 2

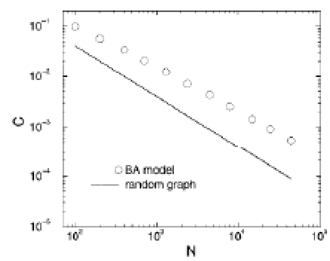


preferential

Additional Properties of the BA Model

- Characteristic Path Length
  - It can be shown (but it is difficult) that the BA model has a length proportional to log(n)/log(log(n))
  - Which is even shorter than in random networks
  - And which is often in accord with – but sometimes underestimates – experimental data
- Clustering
  - There are no analytical results available
  - Simulations shows that in scale-free networks the clustering decreases with the increases of the network order
  - As in random graph, although a bit less
  - This is not in accord with experimental data!





Problems of the BA Model

- The BA model is a nice one, but is not fully satisfactory!
- The BA model does not give satisfactory answers with regard to clustering
  - While the small world model of Watts and Strogatz does!
  - So, there must be something wrong with the model..
- The BA model predicts a fixed exponent of 3 for the power law
  - However, real networks shows exponents between 1 and 3
  - So, there most be something wrong with the model

Reading

- M. E. J. Newman, **Power laws, Pareto distributions and Zipf's law**, Contemporary Physics 46, 323-351 (2005)
- Newman, Mark. **Networks: an introduction**. Oxford University Press, 2010. (Chapter 14)
- Van Steen, Maarten. "**Graph Theory and Complex Networks** An Introduction, 2010. (Chapter 7)
- Barabasi A-L, Albert R (1999) **Emergence of scaling in**
- **random networks**. Science 286: 5009-5012
- Easley, Kleinberg, **Networks, Crowds, and Markets**", 2010, (Chapter 18)
- Lada Adamic, Zipf, Power-laws, and Pareto - a ranking tutorial, http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html