

Power law networks

Social and Technological Networks

Rik Sarkar

University of Edinburgh, 2016.

Piazza forum is up

- Sign up at
 - <https://piazza.com/ed.ac.uk/fall2016/infr11124/>
 - (Link on web page)
- We will use for discussions about lectures, project...
 - Feel free to ask questions
 - Or answer them!
- Note:
 - My knowledge is not perfect! Feel free to point out errors, suggestions, improvements etc..

Degree distribution

- A more sophisticated way of characterizing networks
- More complex than single numbers
- Many standard networks are known to have “standard” degree distributions
- Gives ways to incorporate notions of “popularity” and understand them

Degree distribution

- As a function of k , what fraction of pages in the network have k links?
- A histogram
- What does it look like in a random graph?

Degree distribution of a random graph

- Probability that a node has degree k is:
 - Given by binomial distribution:

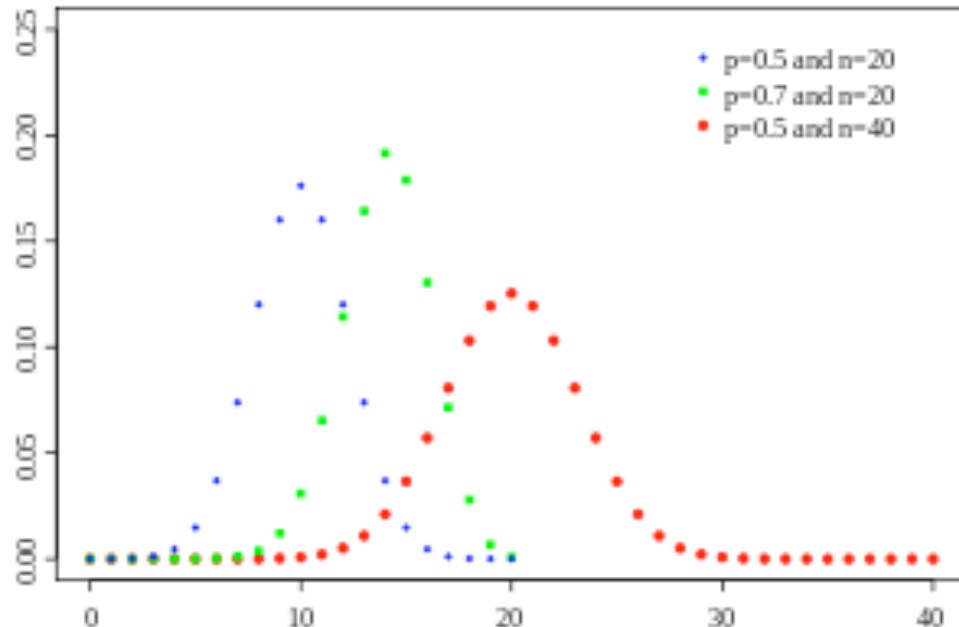
$$\binom{n-1}{k} p^k (1-p)^{n-1-k}$$

Possible sets of
k edges

Probability that
all k are chosen

Probability that
others are not chosen

Degree distribution in a random graph



- Probabilities fall off really fast away from the peak
 - Exponentially fast with k
 - Very low and high degree are very very unlikely

Degree distribution in www

- Suppose we take a real network like the world wide web, and compute degree distribution. What does that look like?
- Let's try.

Degree distribution in www

- For www snapshots, degree distribution follows approximately $\frac{1}{k^2}$



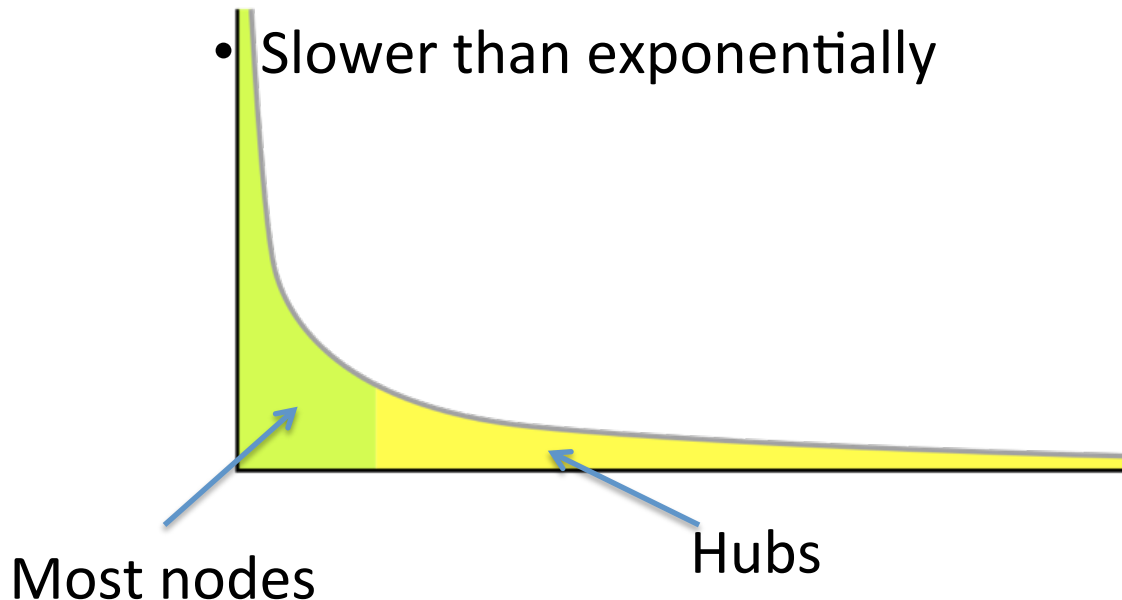
Power law networks

- With degree distribution $\frac{1}{k^\alpha}$
- For some constant α
 - Nothing to with expansion or anything else



What do power law networks mean

- Most nodes have a low degree
- There are several hubs with high degree
 - Heavy tail
 - Probability drops polynomially
- Slower than exponentially



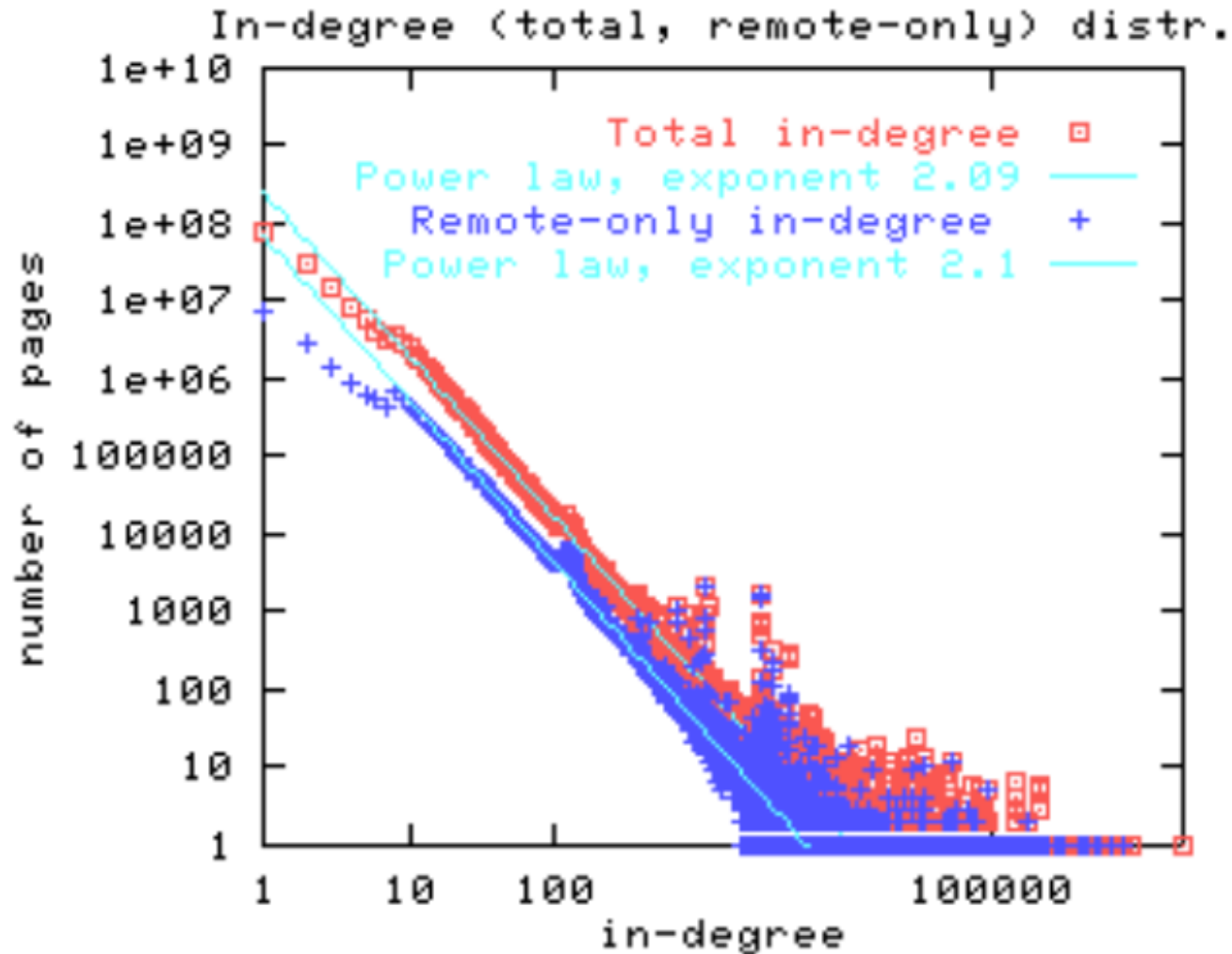
Hubs in power law networks

- Highly connected people/entities
- Critical in information dissemination
- Causes the network to have small diameter
- Examples
 - www, internet..
 - Social networks
 - Collaboration networks

Log log plots

- On ipython notebook

Log log plots for power law are nice and straight



Be careful with log log plots

- The “straight” part needs to extend quite a few orders of magnitude
- Fitting the straight line to determine the right coefficient α is not trivial due to non-linear nature of data
- Beware: log-normal distributions can look similar to power law.

Mean degree in a power law distribution

- The mean is finite iff $\alpha > 2$
 - (On an infinite graph)
- On the www α is slightly larger than 2

Model of power law networks

- We want a model that can be used to create power law networks
- Preferably one that mimics creation of actual power law networks like www
 - Gives us some idea of how these networks were created

Preferential attachment mechanism

- Idea: older and established (popular) sites are likely to have more links to them (yahoo, google...)
- So how about: When a new page arrives, it links to older pages in proportion to their popularity
- When a new link is created on a new page, randomly to older pages with probability of hitting a page x proportional to current popularity of x (number of links to x)

Preferential attachment model

- Takes a parameter p in $[0,1]$
- On a new page, create k links as follows:
- When creating a new link:
 - With probability p
 - Assign it with preferential attachment mechanism
 - With probability $1-p$
 - Assign it with uniform random probability to any existing page

Preferential attachment model

- Takes into consideration that popularity is not the only force behind link creation.
- The randomly assigned links model other reasons for link creation.
- Can be proven to produce power law. see [Kempe lecture notes, 2011]
- Produces same exponent as www for $p \sim 0.9$

Power law often appears in other places

- Popularity of books
- Popularity of people, songs,
- Preferential attachment & power law are often a signature of artificial selection and popularity

Other reasons for power law

- Optimization:
 - Power law found in linguistics (word lengths): most frequent words are short
 - Mandelbrot, Zipf : emerges from need for efficient communication
- Random processes:
 - Press space with probability p , else press a random letter key
 - This will produce a power law distribution of word lengths

