Social Network: Information Cascades and Power-law Distribution

Advanced Computer Networks

Summer Semester 2012

- Examples
  - Choosing the side in a war
    - Side A 70% chance to win
    - Side B 30% chance to win
  - Choosing a restaurant in an unfamiliar town
    - Restaurant A 0 guests
    - Restaurant B 100 guest
    - Your private information: B received good comments in Internet
  - Looking into the sky
  - New products, ideas, …
- Influence of human behaviors and decisions
  - Following the crowd
  - Rich get richer

Information Cascades

Information Cascade

- An information cascade may occur when people make decisions sequentially
  - Later people watching the actions of earlier people, and from these actions inferring something about what the earlier people know.
  - In the restaurant example, when the first diners chose restaurant B, they conveyed information to later diners.
  - A cascade then develops when people abandon their own information in favor of inferences based on earlier people's actions.
- Individuals in a cascade are imitating the behavior of others, but it is from rational inferences of limited information

Milgram et.al. [1969]

- Groups of people (from 1-15 people) stand on a street corner and stare up into the sky
- What happen to the passersby?
  - If one person looking up, very few passersby stopped.
  - If 5 person looking up, more passersby stopped, but most still ignored them
  - If 15 people looking up, 45% of passersby stopped and also stared up into the sky
- Reference: Stanley Milgram, Leonard Bickman, and Lawrence Berkowitz. Note on the drawing power of crowds of dierent size. Journal of Personality and Social Psychology, 13(2):79{82, October 1969.

Basic Ingredients

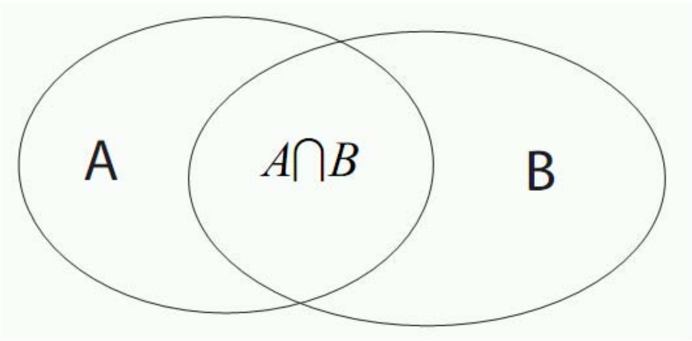
- Herding (Information Cascade)
  - There is a decision to be made
  - People make the decision sequentially
  - Each person has some private information that helps guide the decision
  - You can’t directly observe private information of the others but can see what they do

A Simple Example

- Consider an urn with 3 marbles. It can be either:
  - Majority-blue: 2 blue, 1 red
  - Majority-red: 1 blue, 2 red
- Each person wants to best guess whether the urn is majority-blue or majority-red
- Experiment: One by one each person:
  - Draws a marble
  - Privately looks are the color and puts the marble back
  - Publicly guesses whether the urn is majority-red or majority-blue
- You see all the guesses beforehand.
- How should you make your guess?

- What happens?
  - #1 person: Guess the color you draw from the urn.
  - #2 person: Guess the color you draw from the urn.
    - If same color as 1st, then go with it
    - If different, break the tie by doing with your own color
  - #3 person:
    - If the two before made different guesses, go with your color
    - Else, go with their guess (regardless your color) – cascade starts!
  - #4 person:
    - Suppose the first two guesses BLUE, you go with BLUE (Since 3rd person always guesses BLUE)
  - Everyone else guesses BLUE (regardless of their draw)

Revisit Probabilistic Theory



- Conditional probability of A given B

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}.$$

- Similarly

$$\Pr[B \mid A] = \frac{\Pr[B \cap A]}{\Pr[A]}$$

- Rewriting as: $\Pr[A \mid B] \cdot \Pr[B] = \Pr[A \cap B] = \Pr[B \mid A] \cdot \Pr[A]$
- Bayes' Rule (posterior probability):

$$\Pr[A \mid B] = \frac{\Pr[A] \cdot \Pr[B \mid A]}{\Pr[B]}.$$

Analysis

- #1 follows her own color (private signal)
  - Prior probabilities $\Pr[\textit{majority-blue}] = \Pr[\textit{majority-red}] = \frac{1}{2}.$
  - For the two kinds of urns

$$\Pr[\textit{blue} \mid \textit{majority-blue}] = \Pr[\textit{red} \mid \textit{majority-red}] = \frac{2}{3}.$$

  - If a blue marble is drawn

$$\Pr[\textit{majority-blue} \mid \textit{blue}] = \frac{\Pr[\textit{majority-blue}] \cdot \Pr[\textit{blue} \mid \textit{majority-blue}]}{\Pr[\textit{blue}]}.$$

$$\begin{aligned}\Pr[\textit{blue}] &= \Pr[\textit{majority-blue}] \cdot \Pr[\textit{blue} \mid \textit{majority-blue}] + \\ &\quad \Pr[\textit{majority-red}] \cdot \Pr[\textit{blue} \mid \textit{majority-red}] \\ &= \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}.\end{aligned}$$

  - Thus

$$Pr[majority - blue|blue] = \frac{\frac{1}{2} \cdot \frac{2}{3}}{\frac{1}{2}} = \frac{2}{3}$$

computer NETWORKS

- #2 student
  - #2 knows #1’s color. So, #2 gets 2 colors.
    - If they are the same, decision is easy.
    - If not, break the tie in favor of her own color
- #3 follows majority signal
  - Knows #1, #2 acted on their colors. So, #3 gets 3 signals.
    - If #1 and #2 made opposite decisions, #3 goes with her own color.
    - If #1 and #2 made same choice, #3 follows then (Her decision conveyed no info. Cascade has started!)

$$Pr[majority - blue|blue, blue, red] =?$$

$$Pr[majority - blue|blue, blue, red] = ?$$

- According to Bayes' Rule

$$\Pr[majority\text{-}blue \mid blue,\ blue,\ red] = \frac{\Pr[majority\text{-}blue] \cdot \Pr[blue,\ blue,\ red \mid majority\text{-}blue]}{\Pr[blue,\ blue,\ red]}.$$

- Since

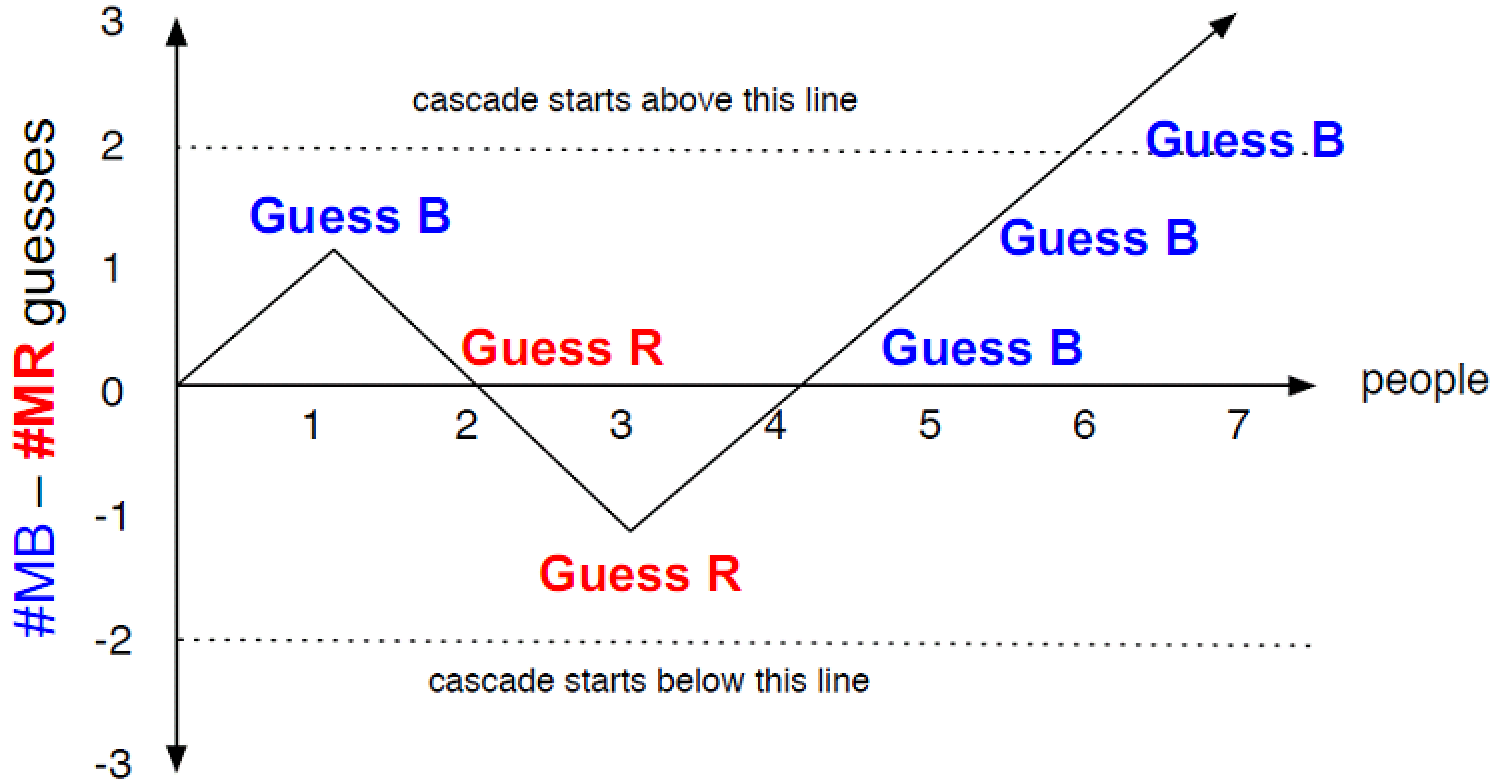
$$\Pr[blue,\ blue,\ red \mid majority\text{-}blue] = \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{4}{27}.$$

$$\begin{aligned}\Pr[blue,\ blue,\ red] &= \Pr[majority\text{-}blue] \cdot \Pr[blue,\ blue,\ red \mid majority\text{-}blue] + \\ &\quad \Pr[majority\text{-}red] \cdot \Pr[blue,\ blue,\ red \mid majority\text{-}red] \\ &= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{6}{54} = \frac{1}{9}.\end{aligned}$$

- So

$$\Pr[majority\text{-}blue \mid blue,\ blue,\ red] = \frac{\frac{4}{27} \cdot \frac{1}{2}}{\frac{1}{9}} = \frac{2}{3}.$$

- Cascade begins when two consecutive students guess the same



Lessons from Cascades

- Cascades can be wrong
  - A cascade of acceptances will start when the first two people happen to get high signals, even though it is the wrong choice for the population
- Cascades can be based on very little information.
  - Since people ignore their private information once a cascade starts, only the pre-cascade information influences the behavior of the population.
- Cascades can be fragile
  - Suppose the first two guess blue
  - Student #1 to #100 guess blue
  - If student #99 and #100 draw red and show then in public
  - Student #101 now has 4 pieces of information, and she guesses based on her own color
  - Cascade is broken!

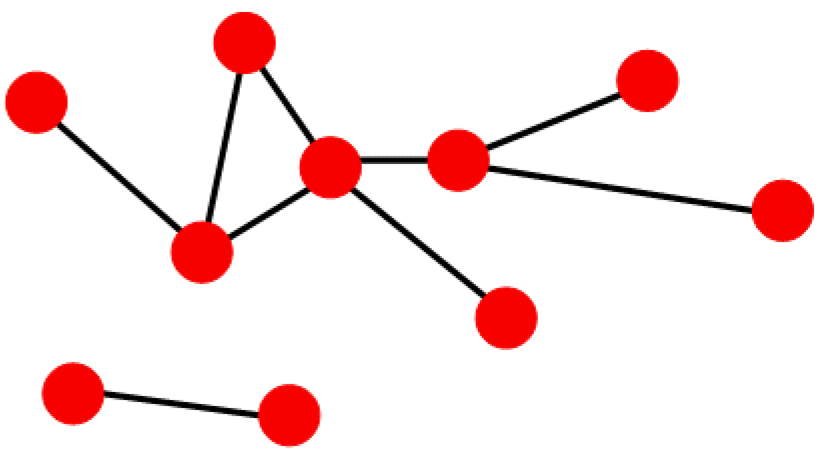
Power-law Distribution

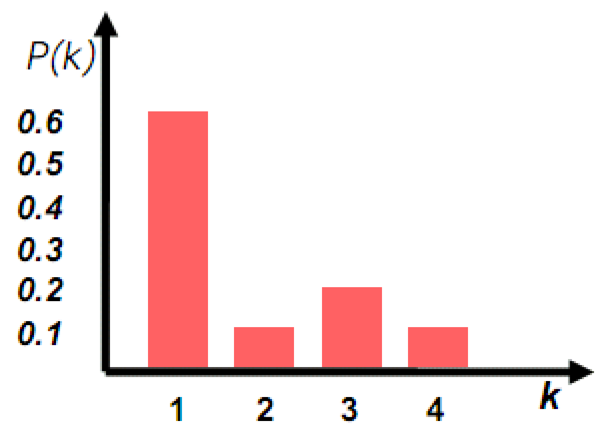
Questions

- How many neighbors does a node have? (degree)
- How far apart are nodes in the network? (distance)
- How close a set of nodes connect with each other? (clustering coefficient)

Degree Distribution

- Degree distribution P(k)
  - Probability that a randomly chosen nodes has degree k
  - N_k: number of nodes with degree k
  - $P(k)=N_k/N$

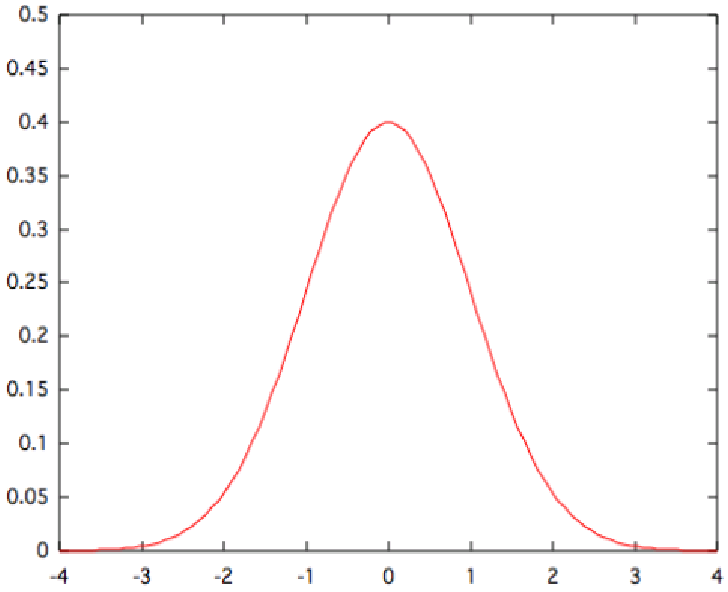




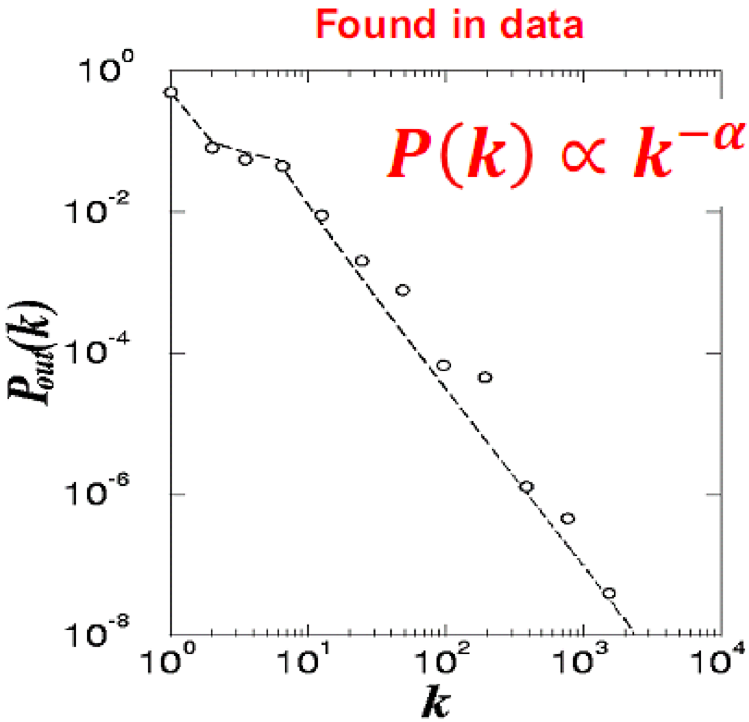
- Popularity of nodes in social networks
  - Imbalance
  - 20% of web pages receives 80% visits
  - Celebrities in Twitter have millions of fans
  - A few rich people own a large amount of wealth
- Node degree distribution
  - What fraction of all nodes have degree k? P(k)=?
  - Normal distribution? – for random graph P(k)=exponential function of -k

Guess: Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Real network: Power-law



- Phenomenon: popularity seems to exhibit extreme imbalances
- Observations:
  - The fraction of telephone numbers that receive k calls per day is roughly proportional to $1/k^2$;
  - The fraction of books that are bought by k people is roughly proportional to $1/k^3$;
  - The fraction of scientific papers that receive k citations in total is roughly proportional to $1/k^3$
  - => The number of monthly downloads for each song at a large on-line music site is proportional to $1/k^c$ for some constant c

- Power-law distribution
  - Let f(k) be the fraction of items have value k

$$f(k) = zk^{-\alpha}$$

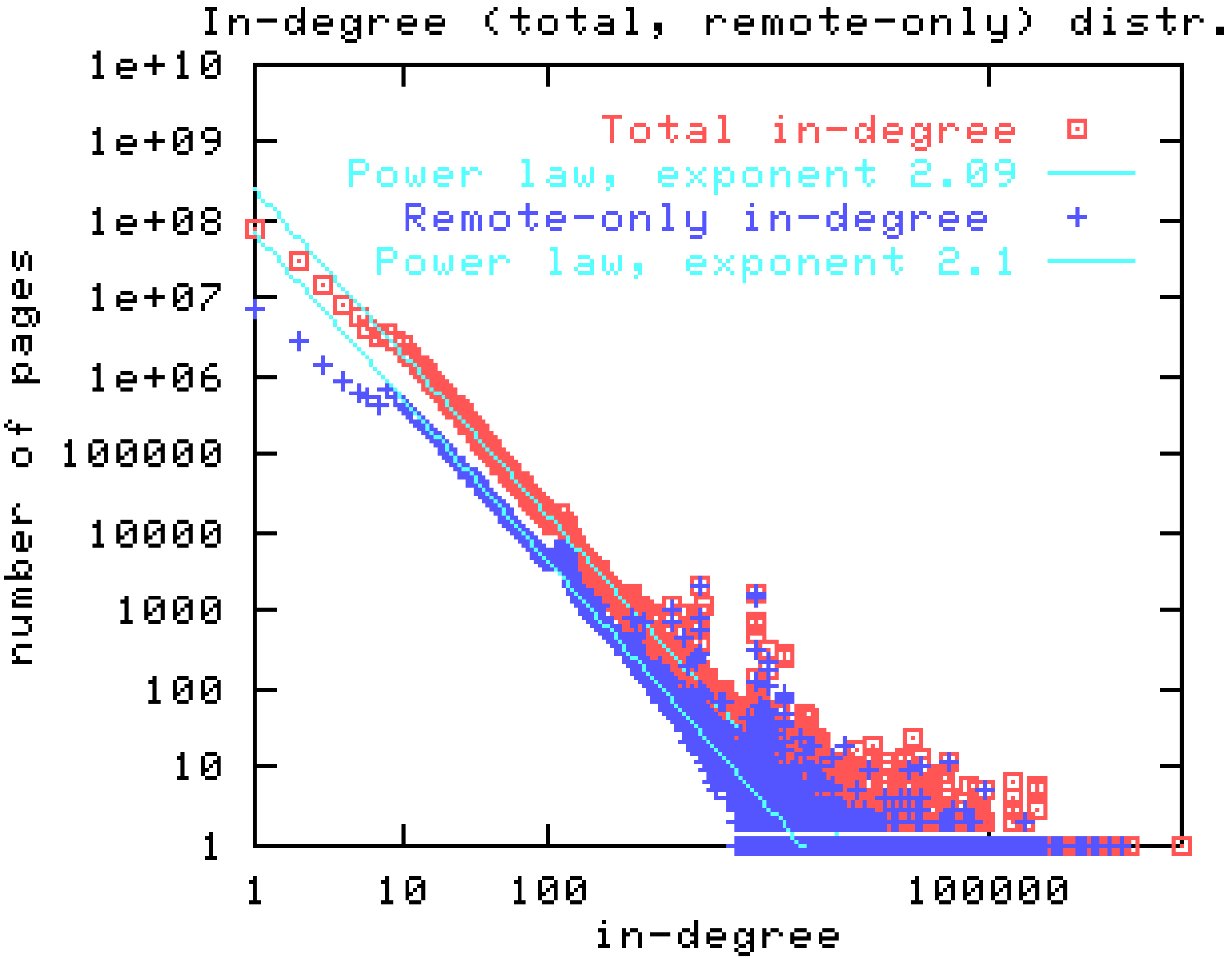
    where α and z are constants

- Taking logarithms of both sides

$$log f(k) = log z - \alpha log k$$

- Testing for power-law distribution
  - If we draw k and f(k) in "log-log" scale, it shows a straight line

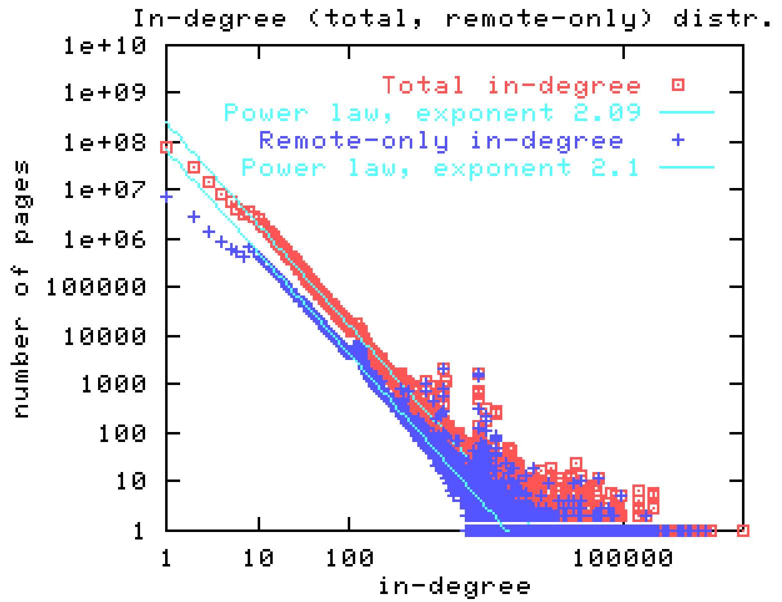
Node Degree of Websites



Estimating power-Law Exponent

- Estimating α ? $$\log f(k) = \log z - \alpha \log k$$
- Simple method: fit a line on log-log axis

$$\min_{\alpha}(\log(y) - \alpha \log(x))^2$$



- For power-law distribution $f(k) = zk^{-\alpha}$
- Estimating the normalizing constant

$$P(x) = z\,x^{-\alpha} \qquad z=?$$

- According to definition

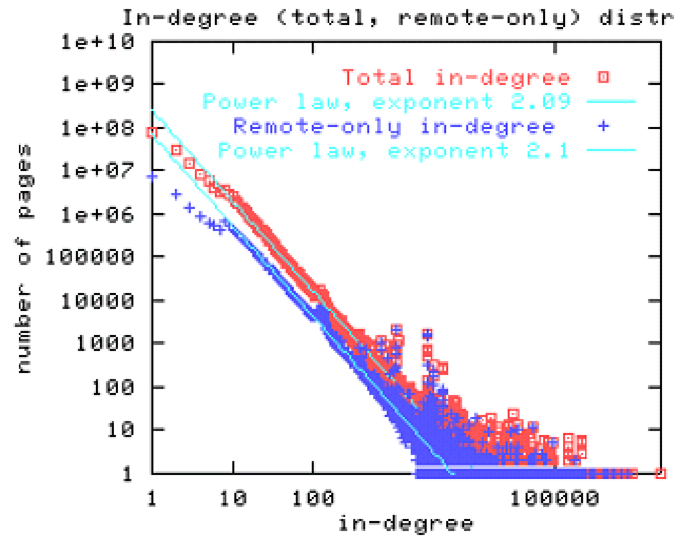
$P(x)$ is a distribution: $\int P(x)dx = 1$

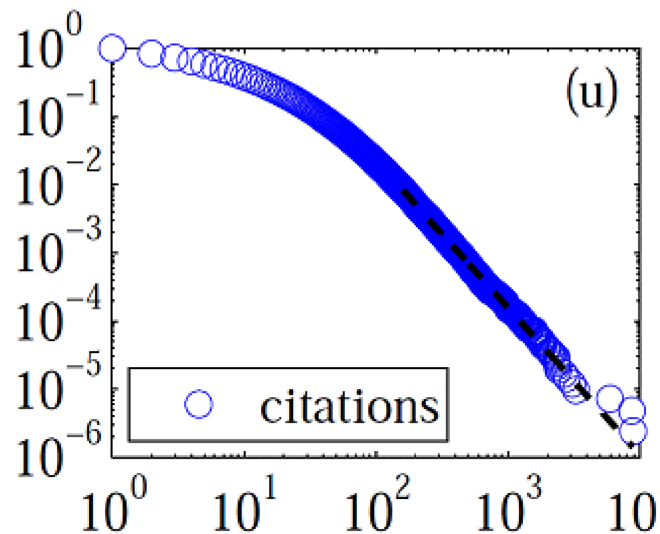
- Thus, let

$$1 = \int_{x_{min}}^{\infty} P(x)dx = z\int_{x_m}^{\infty} x^{-\alpha}dx$$

- *z* can be obtained by solving the above equation

- **Power-law degree exponent is typically $2 < \alpha < 3$**
  - Web graph:
    - $\alpha_{in} = 2.1$, $\alpha_{out} = 2.4$ [Broder et al. 00]
  - Autonomous systems:
    - $\alpha = 2.4$ [Faloutsos^3, 99]
  - Actor-collaborations:
    - $\alpha = 2.3$ [Barabasi-Albert 00]
  - Citations to papers:
    - $\alpha \approx 3$ [Redner 98]
  - Online social networks:
    - $\alpha \approx 2$ [Leskovec et al. 07]





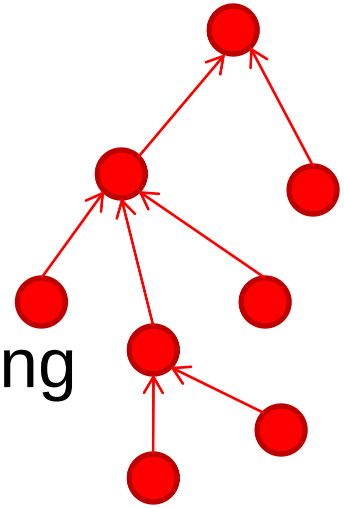
Power-laws Are Everywhere

- Pareto, 1897 – Wealth distribution
- Lotka, Alfred J. 1926 - The frequency of publications by authors in a given field
- Zipf 1940s –Word frequency
- Simon 1950s – City populations
- Holger Ebel 2002- Email Network
- Who-talks-to-whom network
- Number of friends in Facebook
- …

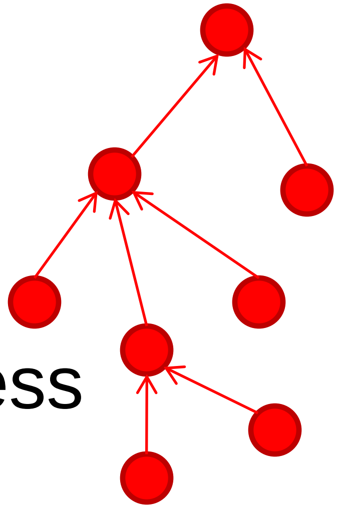
Rich Get Richer Model

- New nodes are more likely to link to nodes that already have high degree
- Power-laws arise from “Rich get richer”
  - We can provide a simple power-law models from consequences of individual decision-making
  - We assume simply that people have a tendency to copy the decisions of people who act before them

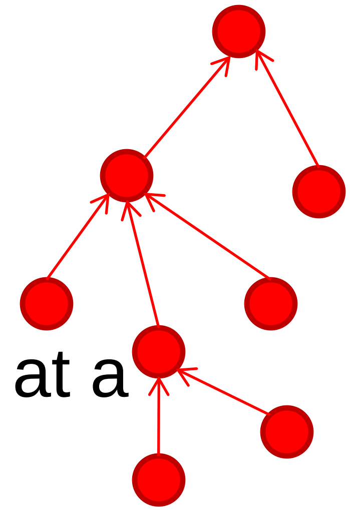
A Simple Model



- A simple model for the creation of links among Web pages
  - Pages are created in order, and named 1; 2; 3; …;N.
  - When page j is created, it produces a link to an earlier Web page i according to:
  - 1) With prob. p (0<p<1), j links to i chosen uniformly at random (from among all earlier nodes)
  - 2) With prob. 1-p, node j chooses node i uniformly at random and links to a node i points to.
    - Note: this is same as saying, with prob. 1-p, node j links to node u with prob. proportional to the degree of u



- The model specifies a randomized process that run for N steps (as the N pages are created one at a time)
- Question: can we determine the expected number of pages with k in-links at the end of the process? (Or analyze the distribution of this quantity?)



- Definition: the number of in-links to a node j at a time step t (t≥j) is a random variable $X_j(t)$
- Facts:
  - Initial condition: Since node j starts with no in-links when it is first created at time j, so $X_j(j) = 0$
  - The expected change to X_j over time
    - In step t+1, node j gains an in-link if and only if the link from the newly created node t+1 points to it
    - With prob. p, node t+1 point to the previous t nodes at random, thus the prob. of linking to node j is p/t
    - With prob. (1-p), node t + 1 links to node j with probability Xj(t)/t
    - So the overall probability that node t + 1 links to node j is

$$\frac{p}{t} + \frac{(1-p)X_j(t)}{t}$$

- in step t+1, the probability of a new link

$$\frac{p}{t} + \frac{(1-p)X_j(t)}{t}$$

- How to calculate $X_j(t)$?
- Approximation
  - Treat it as a deterministic continuous function $x_j(t)$
  - From step t to step t+1 can be viewed as the differential of the function, thus

$$\frac{dx_j}{dt} = \frac{p}{t} + \frac{(1-p)x_j}{t}$$

- Solving the differential equation

$$\frac{dx_j}{dt} = \frac{p}{t} + \frac{(1-p)x_j}{t}$$

  - Let q=1-p, rewriting it as

$$\frac{1}{p+qx_j}\frac{dx_j}{dt} = \frac{1}{t}.$$

  - Integrating both sides

$$\ln(p+qx_j) = q\ln t + c$$

  - Applying the initial condition $x_j(j) = 0$ yields

$$x_j(t) = \frac{1}{q}\left(\frac{p}{j^q}\cdot t^q - p\right) = \frac{p}{q}\left[\left(\frac{t}{j}\right)^q - 1\right]$$

$$x_j(t) = \frac{1}{q}\left(\frac{p}{j^q}\cdot t^q - p\right) = \frac{p}{q}\left[\left(\frac{t}{j}\right)^q - 1\right]$$

- $x_j(t)$ is the degree of node j at time t
- What is the fraction of nodes whose degree ≥k?
  - If a node's degree no smaller than k, it satisfies

    $$x_j(t) = \frac{p}{q}\left[\left(\frac{t}{j}\right)^q - 1\right] \geq k,$$

  - Solving the inequality, we have $j \leq t\left[\frac{q}{p}\cdot k + 1\right]^{-1/q}$
  - Since there are t nodes at time t. the fraction of nodes is:

    $$\frac{1}{t}\cdot t\left[\frac{q}{p}\cdot k + 1\right]^{-1/q} = \left[\frac{q}{p}\cdot k + 1\right]^{-1/q}$$

- What is the fraction of nodes with degree exactly k?
  - Take derivative of the equation, we get $\frac{1}{q}\frac{q}{p}\left[\frac{q}{p}\cdot k + 1\right]^{-1-1/q}$
- So it is a power-law distribution with exponent $1 + \frac{1}{q} = 1 + \frac{1}{1-p}$

computer NETWORKS

Summary

- Information cascades and rich get richer may explain many social behaviors
  - Rumors
  - New technology
  - Fashions
  - Keeping your money or not in a stock market
  - Voting for popular candidates
  - Best selling books, music
  - Riots, protests, strikes