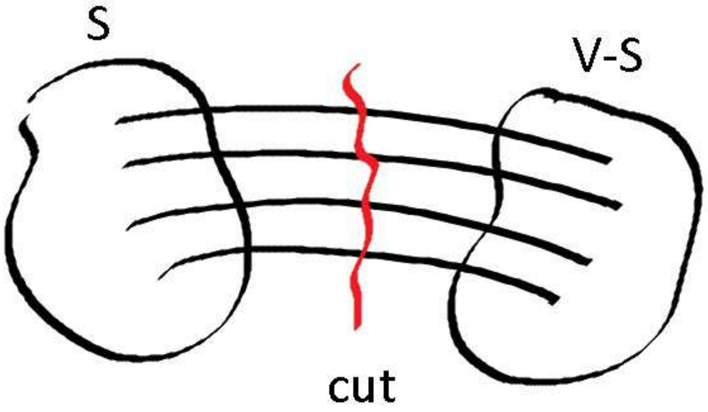
Bisection



$$|S| = |V - S|$$

$$bw(G) = \min_{S \subset V, |S| = |V|/2} \text{cut}(S, V - S)$$

- $bw(G) \geq \frac{n}{4} \lambda_2(L)$

Algorithm for finding sparse cuts: compute Fiedler vector, sort the vertices according to its values , and find the best cut .

Partitioning problem, Walshaw's partitioning archive, DIMACS competition
[BMSSS] "Recent advances in graph partitioning", 2013

p-discrepancy

We define $p-discrepancy$ $\sigma_p(G,\psi)$ of labeling ψ as

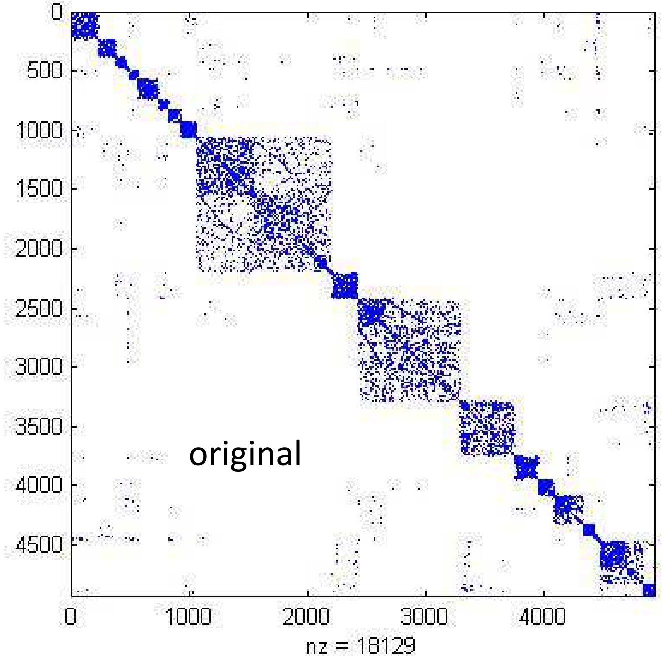
$$\sigma_p(G,\psi) = \left(\sum_{ij\in E} w_{ij}|\psi(i)-\psi(j)|^p\right)^{1/p} \quad \text{and } \sigma_\infty(G,\psi) = \max_{ij\in E} w_{ij}|\psi(i)-\psi(j)|$$

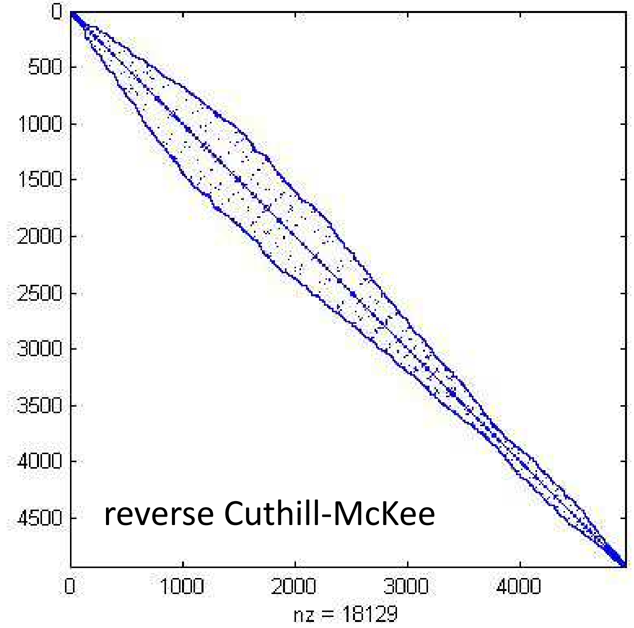
Finding minimum $p-discrepancy$ (such as minimum linear arrangement, 2-sum, bandwidth) is NP-hard.
Basic facts:

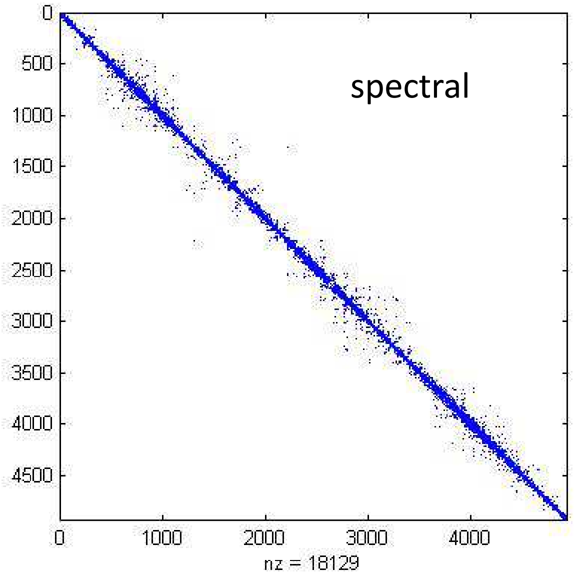
$$\lambda_2 \frac{n(n^2-1)}{12} \le \sigma_2(G,\psi)^2 \le \lambda_n \frac{n(n^2-1)}{12}$$

$$\lambda_2 \frac{(n^2-1)}{6} \le \sigma_1(G,\psi) \le \lambda_n \frac{(n^2-1)}{6}$$

Homework: Juvan, Mohar "Optimal linear labelings and eigenvalues of graphs" (submit by 3/13)


original
nz = 18129


reverse Cuthill-McKee
nz = 18129


spectral
nz = 18129

Matlab demo

Random Walks

Let $p_i(t)$ be the probability that the walk is at i at time t

$$p_i(t) = \sum_j \frac{A_{ij}}{d(j)} p_j(t-1) \text{ or } p(t) = AD^{-1}p(t-1)$$

Another useful form of this relation is

$$D^{-1/2}p(t) = \left(D^{-1/2}AD^{-1/2}\right)\left(D^{-1/2}p(t-1)\right)$$

As $t \to \infty$ the probablity distribution is represented by $p = AD^{-1}p$ or

$$(I - AD^{-1})p = (D - A)D^{-1}p = LD^{-1}p = 0,$$

i.e., $D^{-1}p$ is an eigenvector of L with eval 0.
Example: G is connected $\Rightarrow$ there is only one eval 0 $\Rightarrow$ $D^{-1}p = \alpha 1$, i.e.,

$$p_i = d(i)/\sum_j d(j) = \frac{d(i)}{2m}$$

Intuition: high degree nodes are more likely to be visited.

First Passage Time

Q: what is the mean first passage time the random walk started at *i* reaches *j*?

Absorbing random walk = random walk with one or more nodes we can move to but not leave

We consider an absorbing random walk with **single absorbing vertex *v*.**

Probability that rw has fpt exactly t is $p_v(t) - p_v(t-1)$, i.e., the mean is

$$\tau = \sum_{t=0}^{\infty} t(p_v(t) - p_v(t-1)).$$

For any $i \neq v$ $p_i(t) = \sum_j \frac{A_{ij}}{d(j)} p_j(t-1) = \sum_{j \neq v} \frac{A_{ij}}{d(j)} p_j(t-1)$, i.e.,

$$p'(t) = A'D'^{-1}p'(t-1) = \left(A'D'^{-1}\right)^t p'(0) \quad (' \text{ means } v \text{ is removed})$$

Since $\sum_i p_i(t) = 1$ at all times

$$p_v(t) = 1 - \sum_{i \neq v} p_i(t) = 1 - 1^T p'(t) \text{ and}$$

$$\tau = \sum_{t=0}^{\infty} t 1^T (p'_v(t-1) - p'_v(t-1)) = 1^T \left(I - A'D'^{-1}\right)^{-1} p'(0) = \ldots = 1^T D' L'^{-1} p'(0)$$

Random Models

• Model $G(n,m)$ is a probability distribution $P(G)$ over all graphs with n nodes and m edges.
Properties of model = properties of ensemble
Examples:

- graph diameter $l(G)$ means $\langle l \rangle = \sum_G P(G) l(G) = \frac{1}{\Omega} \sum_G l(G)$
- degree $\langle d(\cdot) \rangle = 2m/n$

• Model $G(n,p)$ - graphs with n nodes and independent probability p for placing an edge between two vertices (aka Erdös-Rényi model).
Properties of model = properties of ensemble where G appears with prob

$$P(G) = p^m (1-p)^{\binom{n}{2}-m}$$

and probablity of drawing a graph with m edges from the ensemble is

$$P(m) = \binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m} \text{ and } \langle m \rangle = \sum_{m=0}^{\binom{n}{2}} m P(m) = \binom{n}{2} p$$

- mean degree $\displaystyle\sum_{m=0}^{\binom{n}{2}} \frac{2m}{n} P(m) = \frac{2}{n}\binom{n}{2} p = (n-1)p = c$

mean degree in a graph with exactly *m* edges

- degree distribution
  - node is connected to a particular k others $q_k = p^k (1-p)^{n-1-k}$
  - node is connected to exacly k others $p_k = \binom{n-1}{k} q_k$
  - in large-scale networks $p = c/(n-1)$ can be very small, i.e.,

$$\ln((1-p)^{n-1-k}) = (n-1-k)\ln(1-c/(n-1)) \approx -(n-1-k)\frac{c}{n-1} \approx -c$$

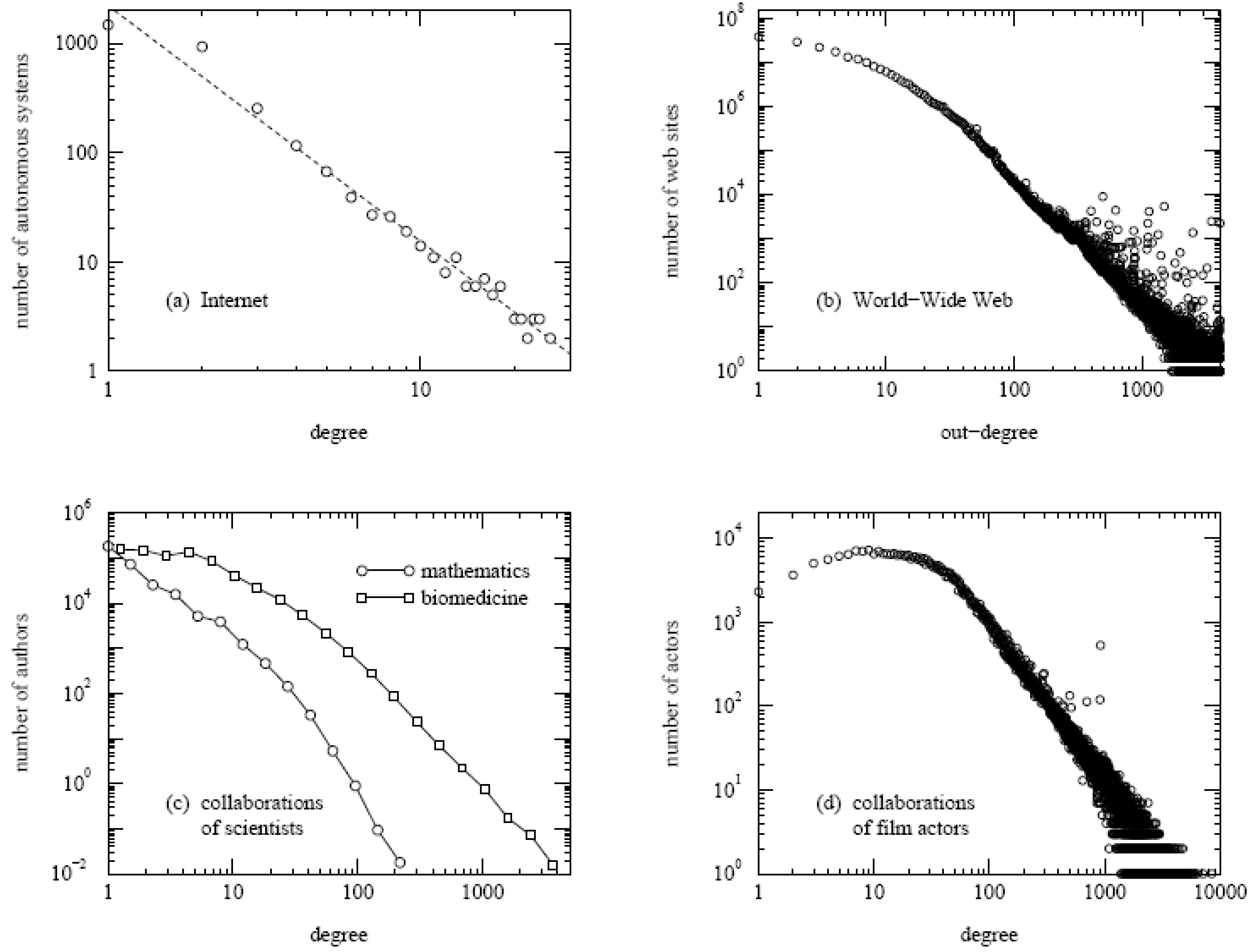
∞

Taylor series reminder: $\ln(1+\frac{1}{x}) = 2\left(A + \frac{1}{3}A^3 + \frac{1}{5}A^5 + \ldots\right)$, where $A = \frac{1}{2x+1}$

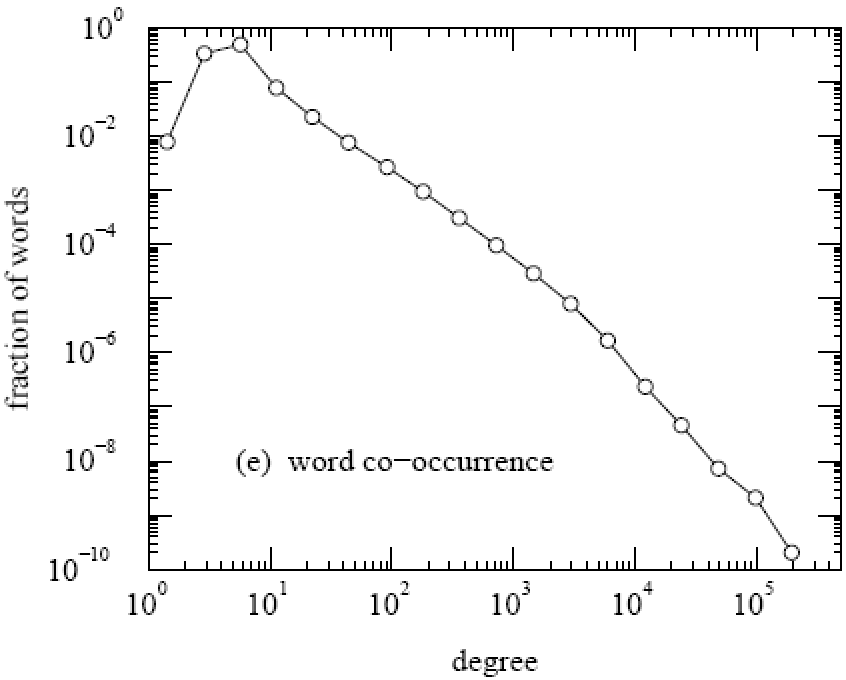
also if $\binom{n-1}{k} = \frac{(n-1)!}{(n-1-k)!k!} \approx \frac{(n-1)^k}{k!}$ then

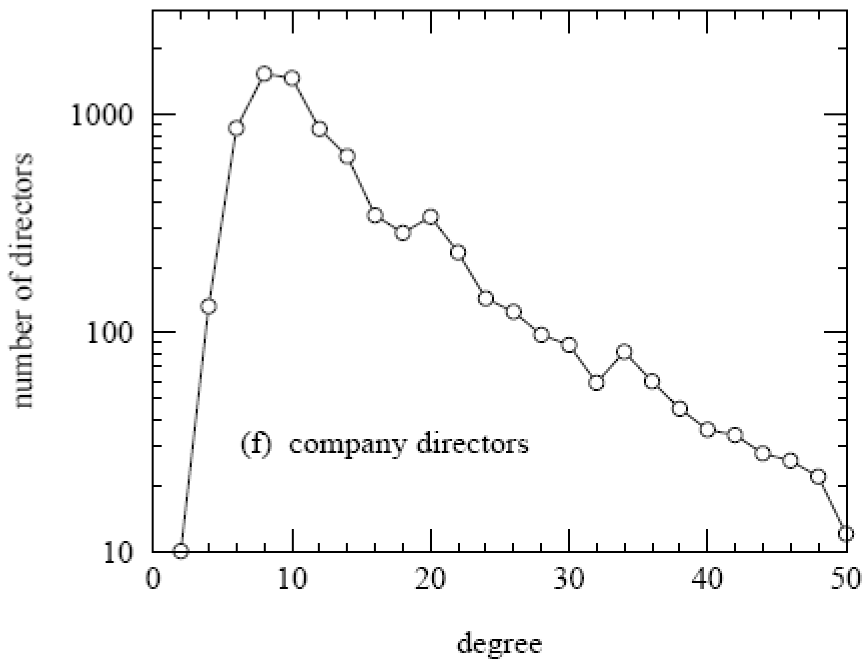
$$p_k = \frac{(n-1)^k}{k!} p^k e^{-c} = \frac{(n-1)^k}{k!}\left(\frac{c}{n-1}\right)^k e^{-c} = e^{-c}\frac{c^k}{k!}$$

In contrast to the degree distribution in random model …



In contrast to the degree distribution in random model …





- clustering coefficient $C = c/(n-1)$ = prob that any two nodes are neighbors

network	n	z	clustering coefficient C measured	clustering coefficient C random graph
Internet (autonomous systems)[a]	6 374	3.8	0.24	0.00060
World-Wide Web (sites)[b]	153 127	35.2	0.11	0.00023
power grid[c]	4 941	2.7	0.080	0.00054
biology collaborations[d]	1 520 251	15.5	0.081	0.000010
mathematics collaborations[e]	253 339	3.9	0.15	0.000015
film actor collaborations[f]	449 913	113.4	0.20	0.00025
company directors[f]	7 673	14.4	0.59	0.0019
word co-occurrence[g]	460 902	70.1	0.44	0.00015
neural network[c]	282	14.0	0.28	0.049
metabolic network[h]	315	28.3	0.59	0.090
food web[i]	134	8.7	0.22	0.065

Newman, “Random graphs as models of networks”