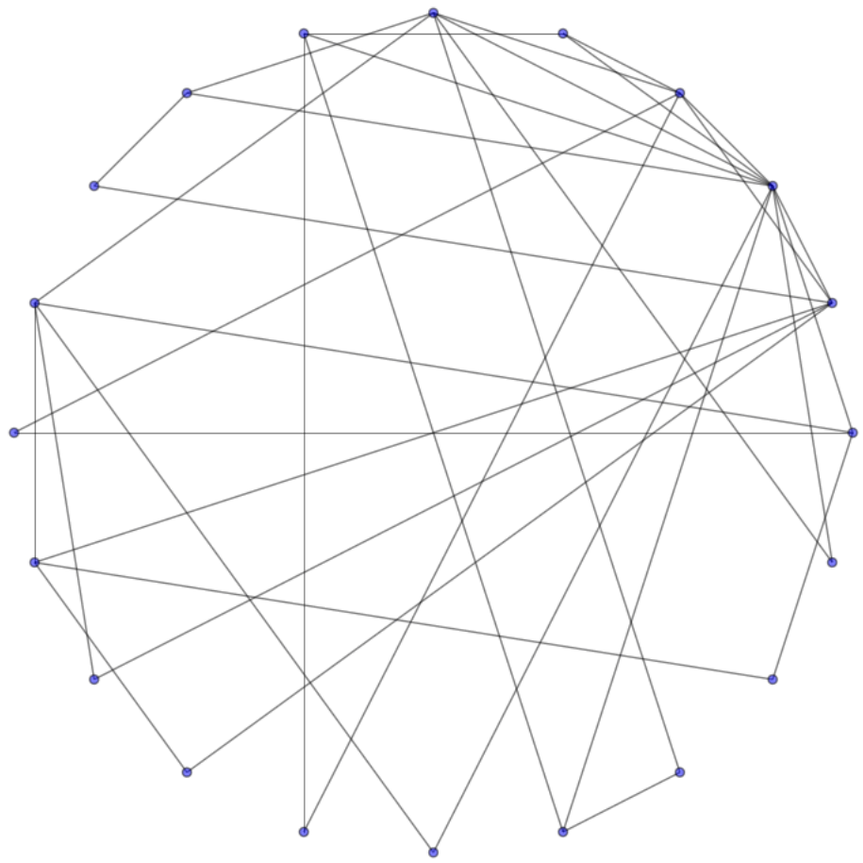
Lecture 9

Review

- Scale-free networks
  (Continuum Theory)
- Simple and elegant
- Capture the process
  → topology as byproduct

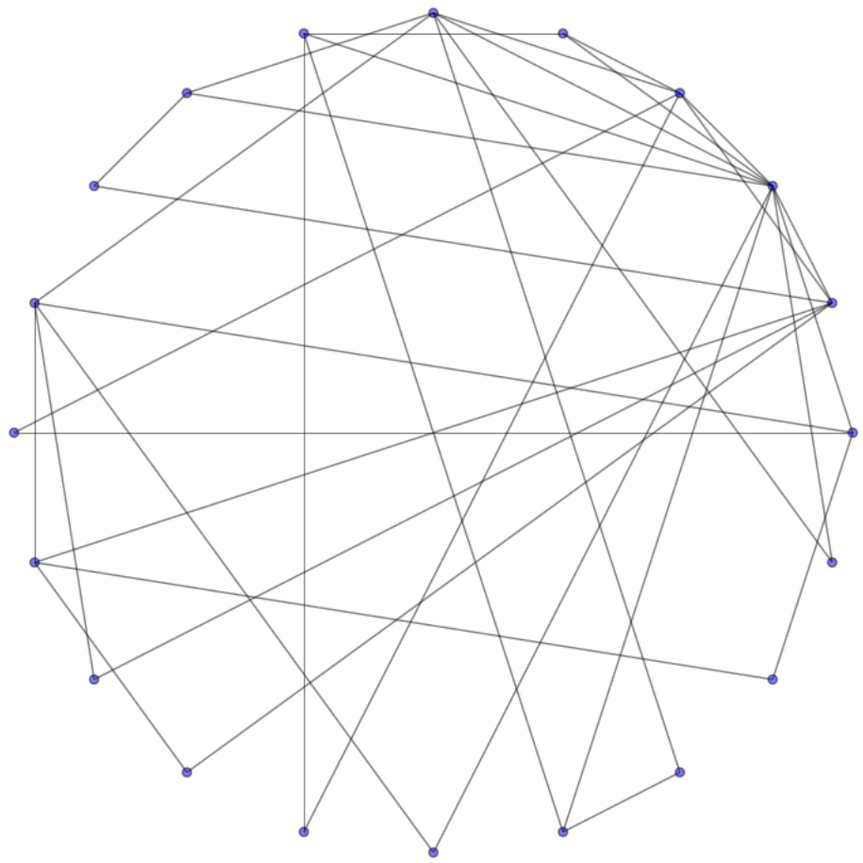
Review

- Scale-free networks (Continuum Theory)
- Simple and elegant
- Capture the process
  → topology as byproduct



Review

- Scale-free networks (Continuum Theory)
- Simple and elegant
- Capture the process
  → topology as byproduct



precise mathematical characterization
v. network generating mechanisms
(why?)

Matthew 25:29

“For to every one that hath shall be given ...

... but from him that hath not, that also which he seemeth to have shall be taken away”

ON A CLASS OF SKEW DISTRIBUTION FUNCTIONS

By HERBERT A. SIMON†

Carnegie Institute of Technology

I. Introduction

It is the purpose of this paper to analyse a class of distribution functions that appears in a wide range of empirical data—particularly data describing sociological, biological and economic phenomena. Its appearance is so frequent, and the phenomena in which it appears so diverse, that one is led to the conjecture that if these phenomena have any property in common it can only be a similarity in the structure of the underlying probability mechanisms. The empirical distributions to which we shall refer specifically are: (A) distributions of words in prose samples by their frequency of occurrence, (B) distributions of scientists by number of papers published, (C) distributions of cities by population, (D) distributions of incomes by size, and (E) distributions of biological genera by number of species.

"the rich get richer"

- Preferential Attachment - proposed mechanism also captures the effect of cumulative advantage

Drawbacks

- All vertices belong to a single component
- Nodes added at constant rate
- Global information
- Undirected network
- Oldest nodes will have the most links (Google?)
- Regard model as directed:
  - attachment proportional to in- + out-degrees?
  - attachment proportional to in-degree?

Questions

- Justify pref. attachment for the in-degree distribution only?
- Acquire links through other mechanism?
- Realistic mechanism?
- Global v. local behavior?
- Relation to node fitness?

  gain function ~ pref. attachment ~ in-degree
- Local gain sensing → local attachment rules

Today

- Class project (example)
- Price's model
- Simon: Wealth distribution
- First to apply Simon idea's in the context of networks
- Master-equation or rate-equation method
- Other growth models (HW)

Example
- The Law -

Assumptions

- Common Law legal system: precedent-based
- Cases grow over time
- Cases get cited by later cases and authorities
- Citation frequency = measure of jurisprudence influence

Assumptions

- Common Law legal system: precedent-based
- Cases grow over time
- Cases get cited by later cases and authorities
- Citation frequency = measure of jurisprudence influence
- Courts invoke the authority of cases

Assumptions

- Common Law legal system: precedent-based
- Cases grow over time
- Cases get cited by later cases and authorities
- Citation frequency = measure of jurisprudence influence
- Courts invoke the authority of cases
- Cases cited approvingly by judges seen as authoritative

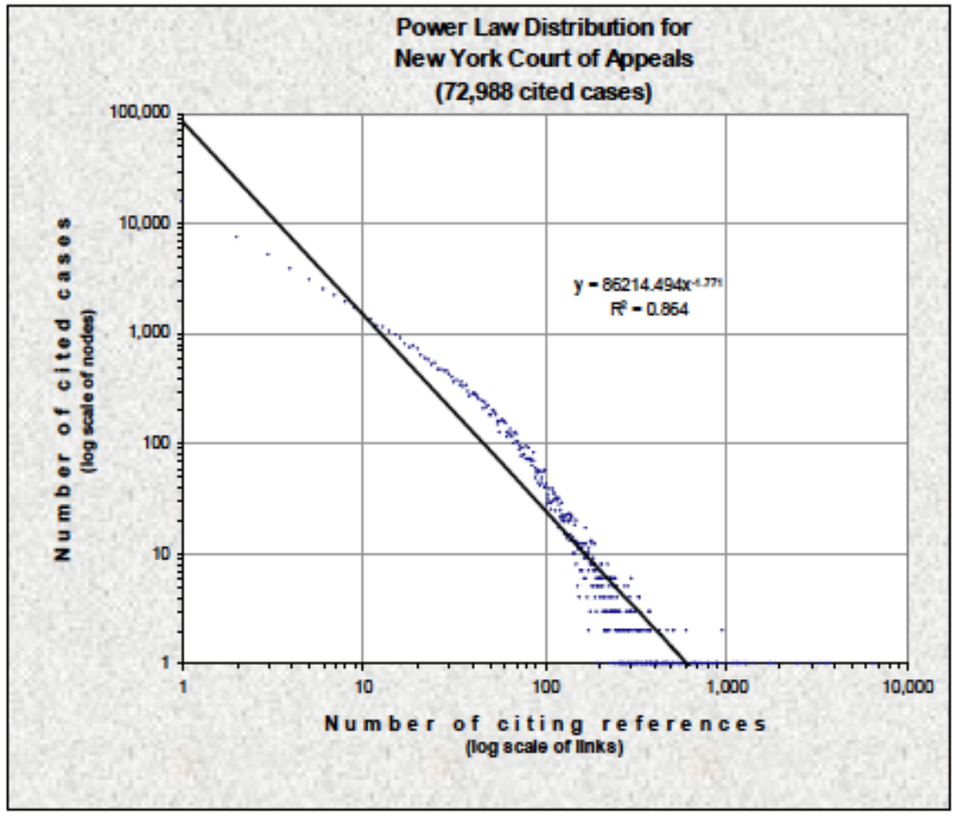
Objectives

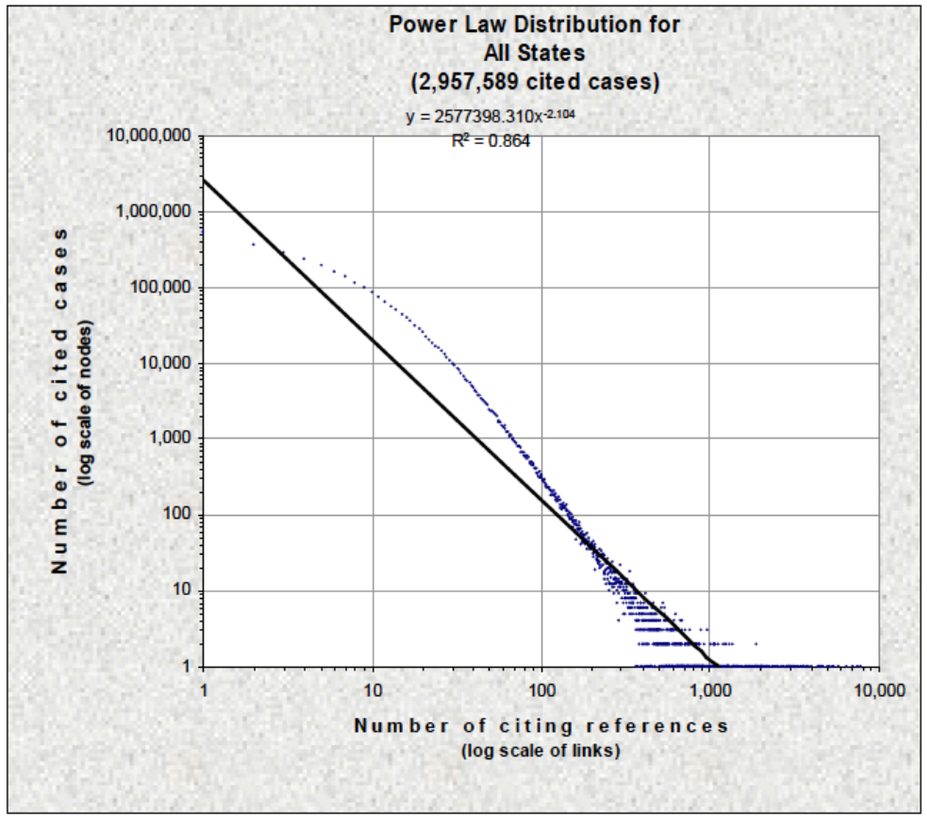
- Understand the law and how it works at a high level of generality
- Understand consequences that follow from the structure of legal citation network

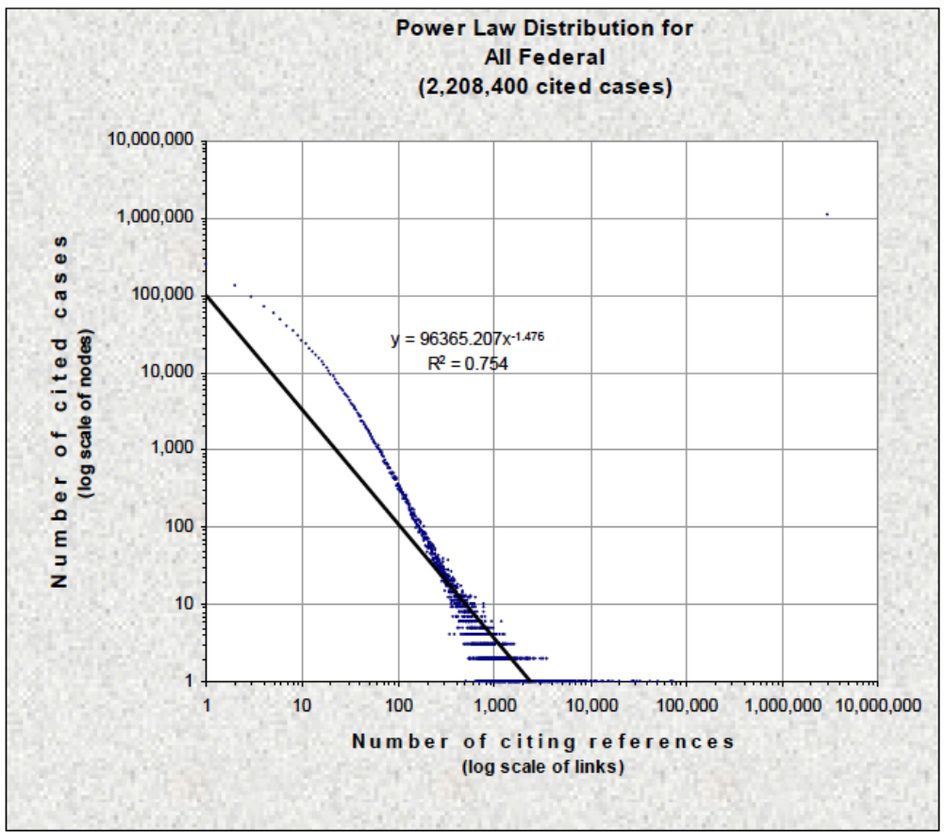
How the Law organizes and shapes?

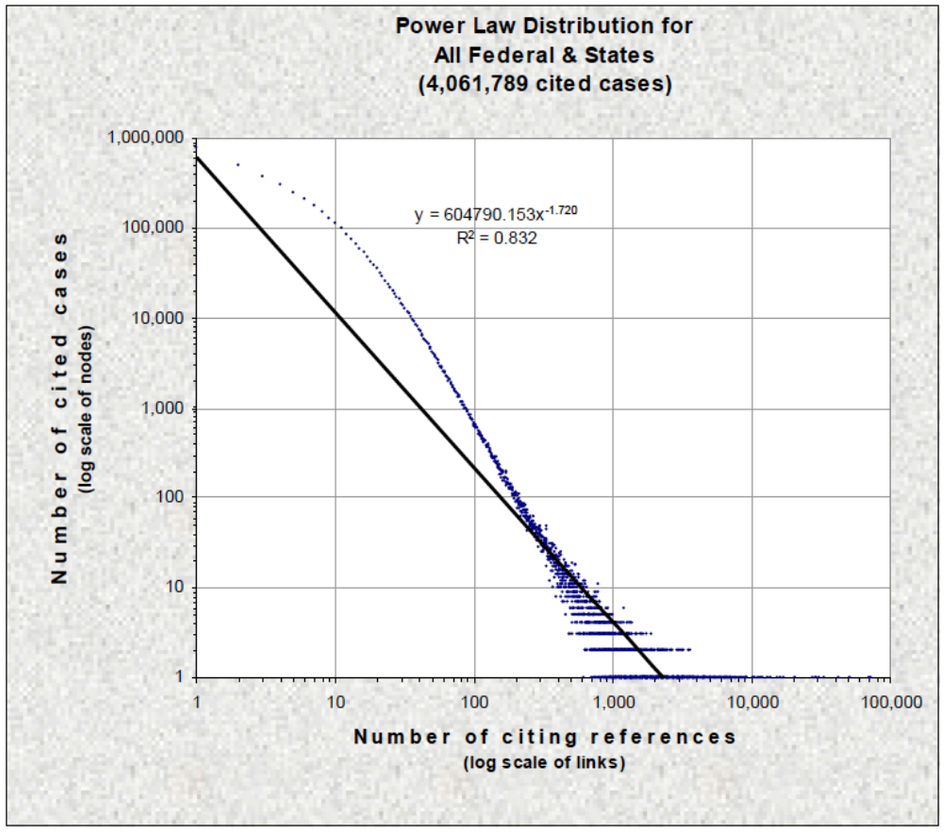
Describe more precisely the consequences of network structure

Citation frequency distribution









Smith, Thomas A., The Web of Law (Spring 2005).
San Diego Legal Studies Research Paper No. 06-11.

Degree distribution

- In each jurisdiction
  - few cases get the majority of all citations
  - most cases are never of rarely cited
- Legal authority concentrated in a relatively small kernel of cases

Small-world

- Multiple sovereigns
- How well integrated?
- Measure of how autonomous, or separate is the state legal system
  - from another
  - from federal system
- Over time as legal profession becomes increasingly specialized?
- International law?

Small-world

- Multiple sovereigns
- How well integrated?
- Measure of how autonomous, or separate is the state legal system
  - from another
  - from federal system
  - Over time as legal profession becomes increasingly specialized?
  - International law?

Understand judicial bodies and the legal regimes they create, in federal and international legal systems

Cluster

- Cases which are tightly linked
- related to each other in terms of meaning and subject matter
- Picture of natural organization of law v. typology

Community

- States form sub-national, supra-state legal regimes? For certain areas of law?
- Certain courts less integrated into the federal judicial system?
- Ninth circuit a “law unto itself?”

Hubs

- Case as a hub
- hub ≠ “most authoritative”
- actively cited v. dead cases
- a decision completely ignored cannot be influential

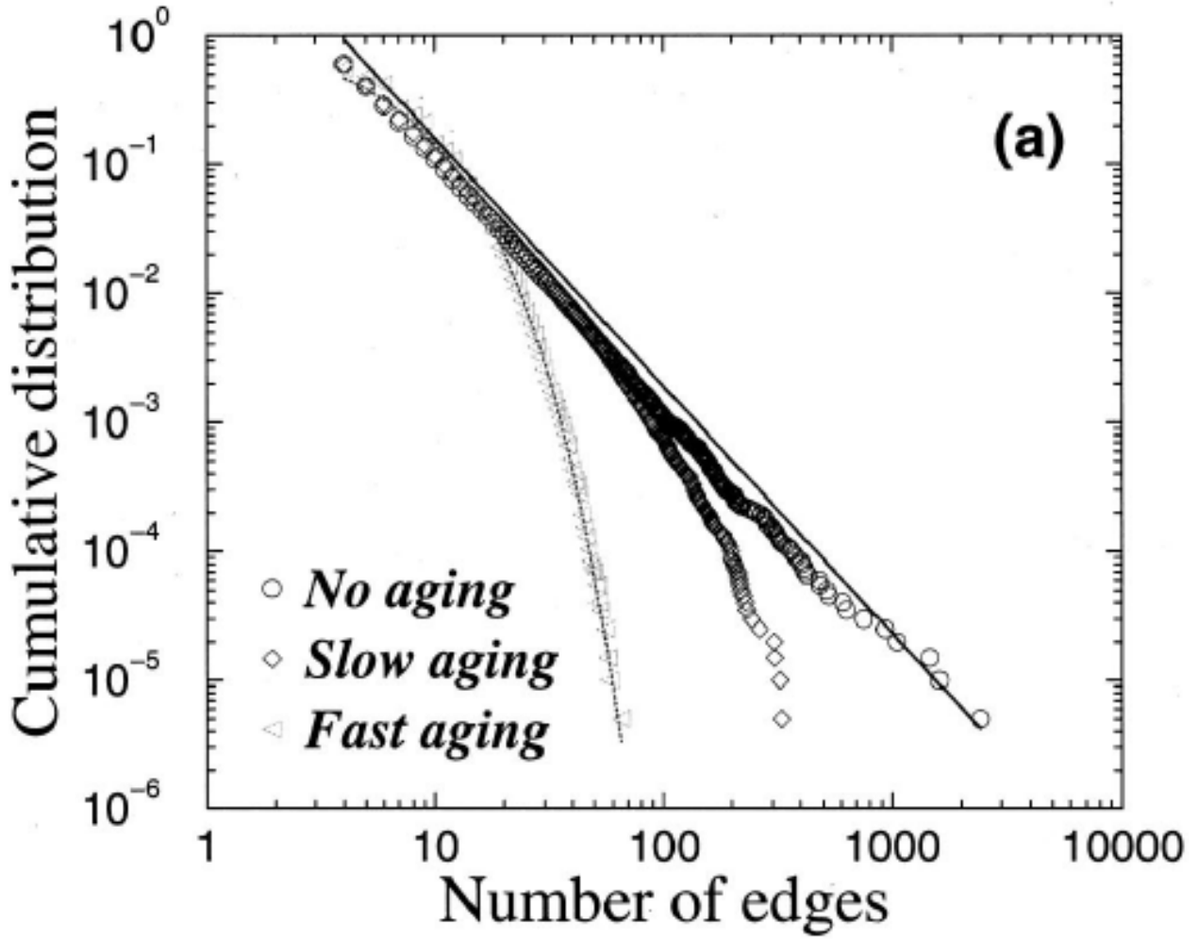
Price's model

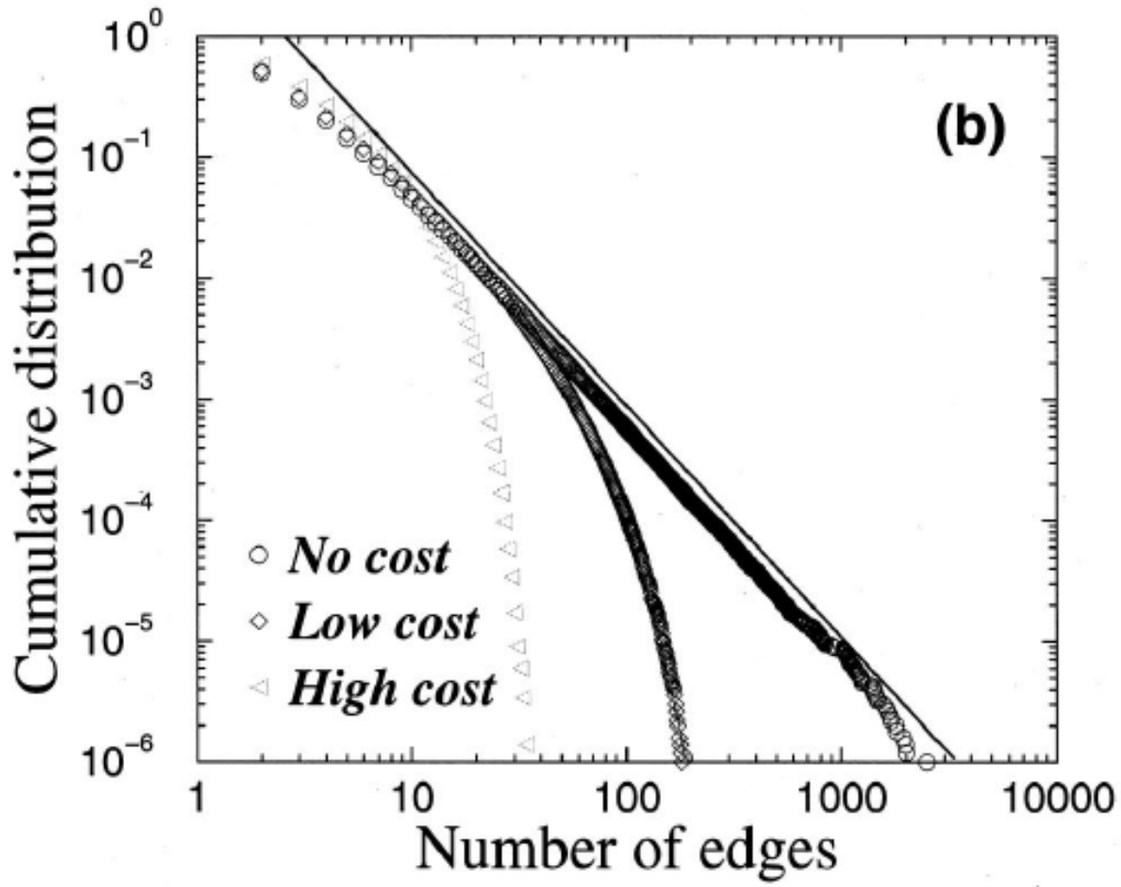
	network	type	n	m	z	ℓ	α	$C^{(1)}$	$C^{(2)}$	r
social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208
	company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88	0.276
	math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34	0.120
	physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.45	0.56	0.363
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60	0.127
	telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1			
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16	
	email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13	0.092
	student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001	−0.029
	sexual contacts	undirected	2 810				3.2			
information	WWW nd.edu	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	−0.067
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7			
	citation network	directed	783 339	6 716 198	8.57		3.0/–			
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	–	0.13	0.15	0.157
	word co-occurrence	undirected	460 902	17 000 000	70.13		2.7		0.44	
technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39	−0.189
	power grid	undirected	4 941	6 594	2.67	18.99	–	0.10	0.080	−0.003
	train routes	undirected	587	19 603	66.79	2.16	–		0.69	−0.033
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082	−0.016
	software classes	directed	1 377	2 213	1.61	1.51	–	0.033	0.012	−0.119
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030	−0.154
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011	−0.366

Homework

Growth constraints

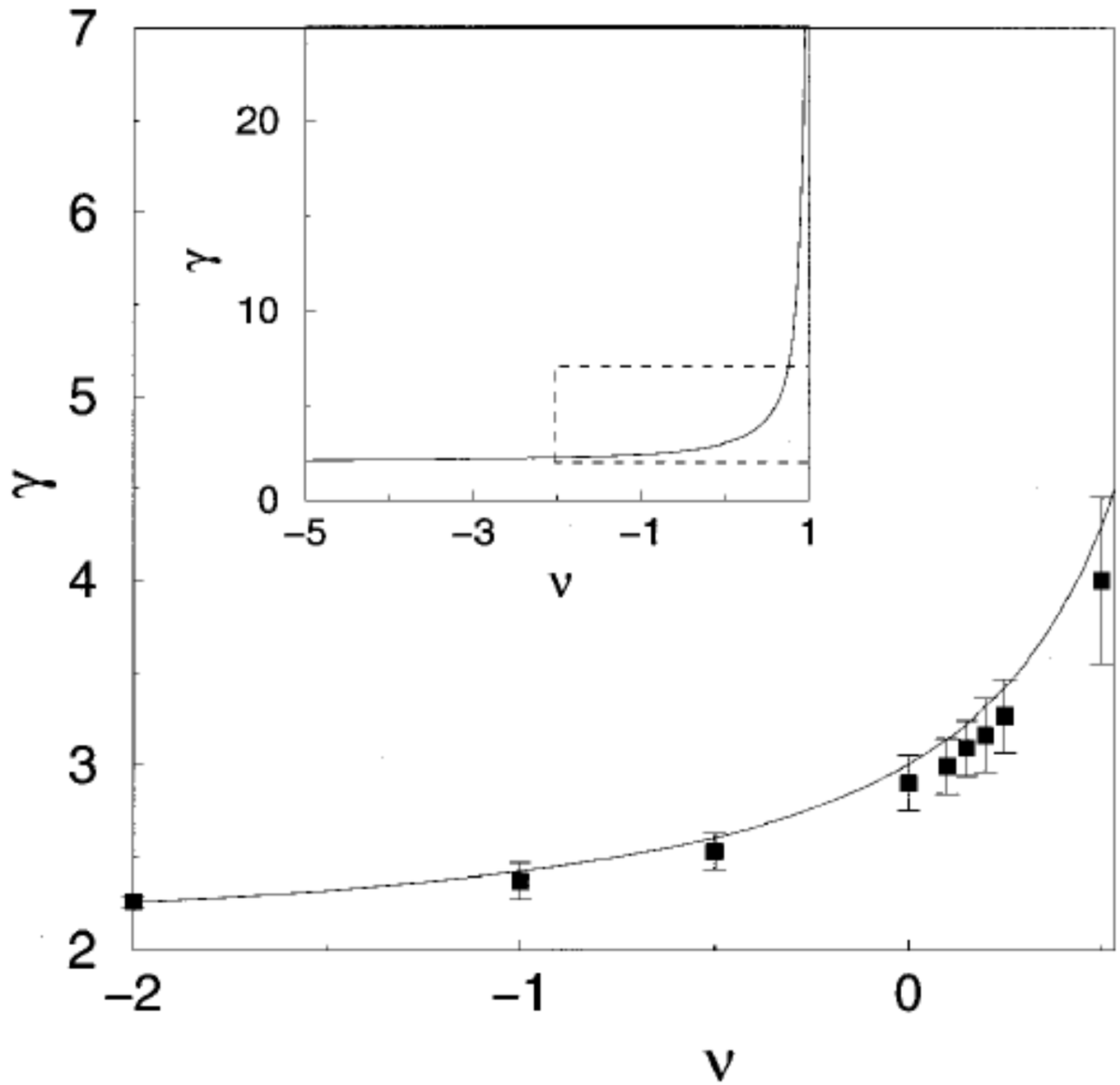
- Finite lifetime (social networks)
- Finite edge capacity
- Evolve following growth + pref. attachment
- No new edges are connected after a certain age or critical number of edges


(a)
Cumulative distribution
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
No aging
Slow aging
Fast aging
1
10
100
1000
10000
Number of edges


(b)
Cumulative distribution
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
No cost
Low cost
High cost
1
10
100
1000
10000
Number of edges

Gradual aging

- Probability that a new nodes connects to a node decays as $(t-t_i)^{-v}$
- The parameter v is a tunable parameter
- Degree distribution depends on v


7
6
5
4
3
2
γ
−2
−1
0
ν
20
10
0
γ
−5
−3
−1
1
ν

Fitness model

- Competitive aspect
- Intrinsic ability to compete for edges (e.g., Google)
- New node with a fitness η_i chosen from a distribution p_η

Remarks

- Patters + statistical regularities on networks
- Motivated by empirical studies or real networks
- These properties the only important quantities to measure?
- Models of networks (2 perspectives)
- Understanding of topology
- Act as a substrate

Next...

- Processes that take pace on networks