

RETI CASUALI

Modello dei grafi casuali

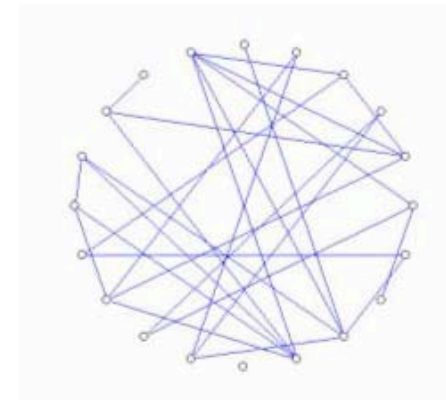
- Un grafo casuale è un grafo con N nodi connessi da n archi, scelti casualmente tra tutte le possibili combinazioni di collegamenti



$p = 0$
(a)



$p = 0.1$
(b)



$p = 0.2$
(c)

On Random Graphs, Erdos and Renyi, 1959.

Combinazioni

- ☉ Da un insieme di oggetti, le combinazioni sono le possibile associazioni di questi oggetti in gruppi di n membri
- ☉ Se volessimo combinare l'insieme ABCDE in tutti i possibili gruppi di 3:

ABC ABD ABE ACD ACE ADE

BCD BCE BDE

CDE

Combinazioni

- Da un insieme di oggetti, le combinazioni sono le possibile associazioni di questi oggetti in gruppi di n membri
- Se volessimo combinare l'insieme ABCDE in tutti i possibili gruppi di 3:

ABC ABD ABE ACD ACE ADE

BCD BCE BDE

CDE

Combinazioni

- Da un insieme di oggetti, le combinazioni sono le possibile associazioni di questi oggetti in gruppi di n membri
- Se volessimo combinare l'insieme ABCDE in tutti i possibili gruppi di 3:

$$5 \times 4 \times 3 / 1 \times 2 \times 3 = 10$$

$$5 \times 4 \times 3 / 3! = 10$$

Combinazioni

- ☉ Da un insieme di oggetti, le combinazioni sono le possibile associazioni di questi oggetti in gruppi di n membri
- ☉ Se volessimo combinare l'insieme ABCDE in tutti i possibili gruppi di 3:

$${}_5C_3 = 10$$

$$C_3^5 = 10$$

Proprietà grafi casuali

- Dato un grafo con N nodi, il numero di archi possibili equivale a $N(N-1)/2$
- Dati N nodi e n archi, Il numero di grafi possibili equivale a $\binom{N(N-1)/2}{n}$. Ad esempio $\binom{10}{5} = 252$
- La costruzione di un grafo casuale può essere fatta semplicemente identificando N nodi e collegandoli tra di loro con una probabilità $p \rightarrow [0, 1]$
 - Quindi il numero medio di archi sarà dato da $n = \binom{N(N-1)/2}{1} p$

Proprietà grafi casuali

- Erdos e Renyi studiarono diverse proprietà dei grafi notando come queste non emergano gradualmente ma non appena venga superata un certa soglia di p
- Studiarono quindi la p critica per cui il grafo acquisisce una proprietà q ; una proprietà importante fu la topologia
- Se il numero di archi inserito era basso il grafo era costruito da diverse componenti non connesse tra loro e connettenti $O(\log N)$ nodi
- Ma per $n = N/2$ i cluster iniziano a connettersi e si ottiene un cluster gigante

Proprietà grafi casuali

- **Grado medio di un grafo $\langle k \rangle$**
- Sarà dato dalla somma di tutti i gradi diviso per il numero di nodi
 - Somma di tutti i gradi è $[N(N-1)/2]*2$ visto che conteremo un arco due volte
 - Moltiplicato per p se il grafo è stato generato da un processo random
 - Quindi $N(N-1)p/N = N-1p$ approssimato a Np per N grande

Proprietà grafi casuali

- **Coefficiente di clustering C_i**
- Sarà dato dal numero n di archi esistenti tra i k_i nodi adiacenti a i e il numero massimo di archi possibili tra essi
 - Quindi $2k_i(k_i-1)p / 2k_i(k_i-1)$ che si semplifica in p
- Il coefficiente di clustering medio può essere calcolato o come media dei vari C_i , e sarà ancora uguale a p
- Oppure come somma di tutti gli archi esistenti fratto il numero massimo di archi possibili: $\sum n_i / \sum \langle k_i \rangle C_2$ che invece dipende da N

Proprietà grafi casuali

- **Distribuzione dei gradi $P(k)$**
- La distribuzione del grado k_i è regolata da una legge binomiale con parametri $N-1$ e p

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

con $k \rightarrow [0, N-1]$

- Segue una distribuzione poissoniana con apice in $\langle k \rangle$

Proprietà grafi casuali

⊗ **Diametro d**

- ⊗ Il diametro di un grafo è la distanza massima tra ogni coppia dei suoi nodi
- ⊗ Data una rete con un grado medio dei nodi pari a k , il numero dei nodi a distanza d da un dato nodo *sarà circa*
 $\langle k \rangle^d$
- ⊗ Si può anche osservare il seguente rapporto: $d = \ln(N) / \ln(\langle k \rangle)$
 - ⊗ all'aumentare di N il diametro cresce lentamente

Proprietà grafi casuali

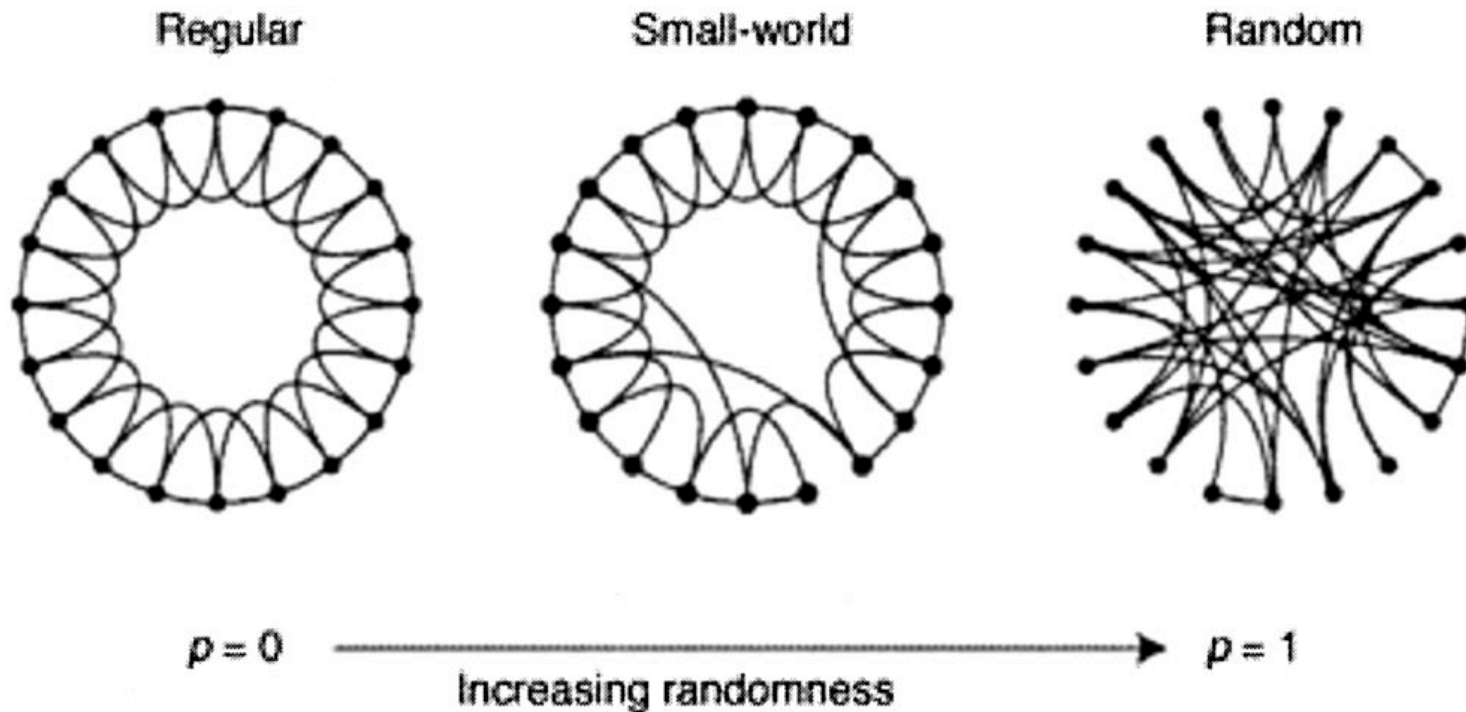
- **Cammino medio tra i nodi L**

- Data dalla somma delle distanze minime $d_{i,j}$ fratto il numero massimo di archi

$$L = 2 \sum d_{i,j} / N(N-1)$$

- Si usa la media armonica per non considerare i nodi non connessi

Modello del "piccolo mondo"



Collective Dynamics of Small World Networks, Watts and Strogatz, 1998.

Proprietà "piccolo mondo"

- ④ **Cammino medio tra i nodi L**

- ④ Per p piccole, $L(p)$ varia linearmente con la dimensione del sistema, avvicinandosi al comportamento di una rete regolare

$$L = 2 \sum d_{i,j} / N(N-1)$$

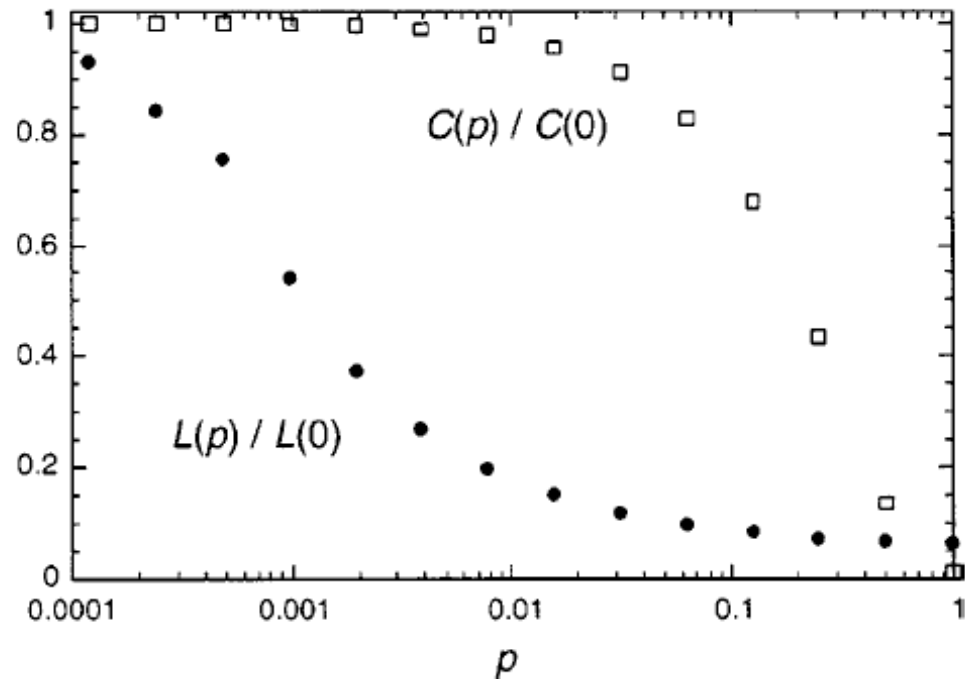
- ④ Mentre per p grandi il comportamento sarà simile a quello di una rete casuale e la variazione è quindi logaritmica

Proprietà "piccolo mondo"

- **Coefficiente di clustering C**
- Sarà dato dal numero di archi esistenti tra i k_i nodi adiacenti a i (m_i) e il numero massimo di archi possibili tra essi
- Quindi $2 m_i / k_i(k_i - 1)$; per piccoli valori di p sarà simile a quello della rete regolare $m_i / k_i(k_i - 1)$; per valori alti di p sarà simile a quello di una rete casuale e cioè p

Proprietà "piccolo mondo"

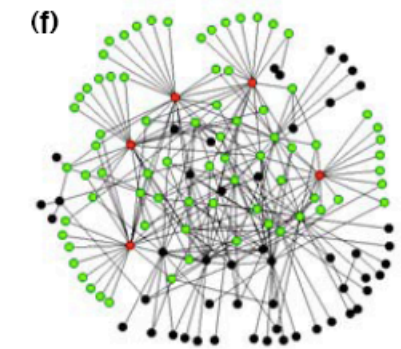
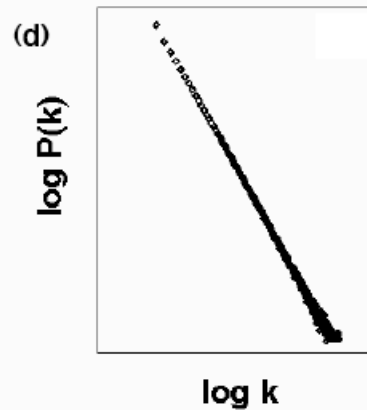
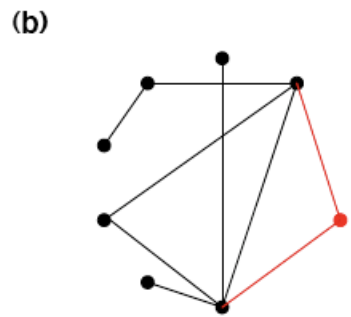
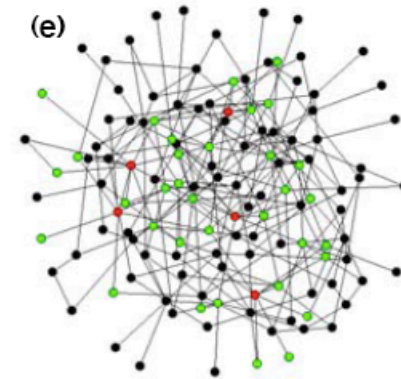
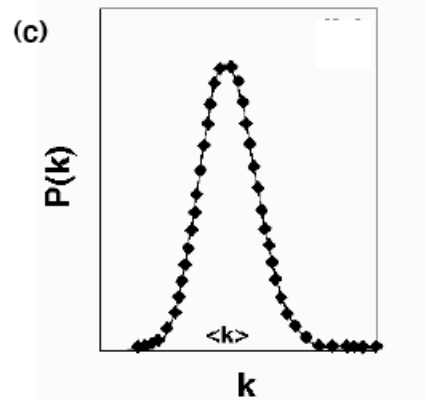
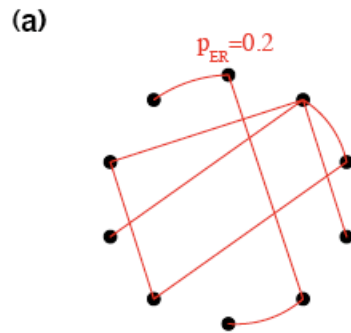
- Questo dimostra l'esistenza di una rete i cui nodi con **C** alto e **L** basso



Proprietà "piccolo mondo"

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference
WWW, site level, undir.	153 127	35.21	3.1	3.35	0.1078	0.00023	Adamic, 1999
Internet, domain level	3015–6209	3.52–4.11	3.7–3.76	6.36–6.18	0.18–0.3	0.001	Yook <i>et al.</i> , 2001a, Pastor-Satorras <i>et al.</i> , 2001
Movie actors	225 226	61	3.65	2.99	0.79	0.00027	Watts and Strogatz, 1998
LANL co-authorship	52 909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman, 2001a, 2001b, 2001c
MEDLINE co-authorship	1 520 251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman, 2001a, 2001b, 2001c
SPIRES co-authorship	56 627	173	4.0	2.12	0.726	0.003	Newman, 2001a, 2001b, 2001c
NCSTRL co-authorship	11 994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman, 2001a, 2001b, 2001c
Math. co-authorship	70 975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási <i>et al.</i> , 2001
Neurosci. co-authorship	209 293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási <i>et al.</i> , 2001
<i>E. coli</i> , substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner and Fell, 2000
<i>E. coli</i> , reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner and Fell, 2000
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya and Solé, 2000
Silwood Park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya and Solé, 2000
Words, co-occurrence	460.902	70.13	2.67	3.03	0.437	0.0001	Ferrer i Cancho and Solé, 2001
Words, synonyms	22 311	13.48	4.5	3.84	0.7	0.0006	Yook <i>et al.</i> , 2001b
Power grid	4941	2.67	18.7	12.4	0.08	0.005	Watts and Strogatz, 1998
<i>C. Elegans</i>	282	14	2.65	2.25	0.28	0.05	Watts and Strogatz, 1998

Modello a invarianza di scala



Barabasi, and Albert, Emergence of scaling in random networks, 1999.

Leggi di potenza

- ◉ Nella distribuzione di Poisson le “ali” della campana decrescono con velocità esponenziale, di conseguenza abbiamo l’assenza di nodi altamente connessi
- ◉ Questo significa che molti nodi con poche connessioni coesistono con pochi nodi altamente connessi
- ◉ In formula, Barabasi ha stabilito empiricamente che la distribuzione dei gradi $P(k)$ decade esponenzialmente all’aumentare del grado k

$$P(k) \rightarrow k^{-\gamma}$$

Leggi di potenza

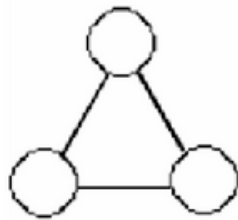
- Una topologia modellata in base ad una legge di potenza non può essere spiegata coi modelli per reti casuali e “piccolo mondo”
 - La costruzione non è dinamica
 - La creazione di collegamenti è casuale

Dinamica di crescita

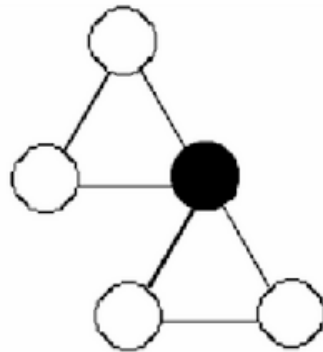
- Barabasi e Albert propongono un modello per descrivere la dinamica di evoluzione della rete, di modo che la sua topologia sia conforme con la legge di potenza
- Introducono quindi la nozione di **collegamento preferenziale**
 - La probabilità Π che un nuovo nodo j si connetta al nodo i , è proporzionale al grado k_i di i

$$\Pi(k_i) = k_i / \sum k_j$$

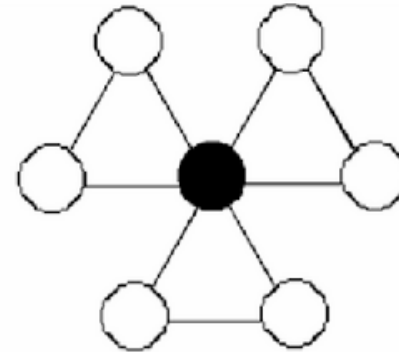
Dinamica di crescita



t=0



t=1

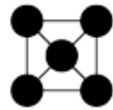


t=2

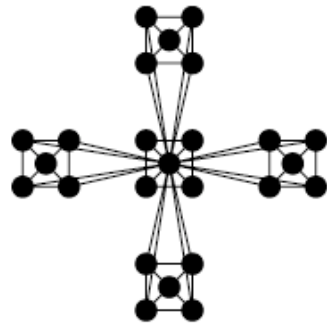
Reti con legge di potenza

Network	Size	$\langle k \rangle$	γ_{out}	γ_{in}
WWW	325,729	4.51	2.45	2.1
WWW	4×10^7	7	2.38	2.1
WWW	2×10^8	7.5	2.72	2.1
WWW, site	260,000			1.94
Internet, domain*	3,015 - 4,389	3.42 - 3.76	2.1 - 2.2	2.1 - 2.2
Internet, router*	3,888	2.57	2.48	2.48
Internet, router*	150,000	2.66	2.4	2.4
Movie actors*	212,250	28.78	2.3	2.3
Coauthors, SPIRES*	56,627	173	1.2	1.2
Coauthors, neuro.*	209,293	11.54	2.1	2.1
Coauthors, math*	70,975	3.9	2.5	2.5
Sexual contacts*	2810		3.4	3.4
Metabolic, E. coli	778	7.4	2.2	2.2
Protein, S. cerev.*	1870	2.39	2.4	2.4
Ythan estuary*	134	8.7	1.05	1.05
Silwood park*	154	4.75	1.13	1.13
Citation	783,339	8.57		3
Phone-call	53×10^6	3.16	2.1	2.1
Words, concurrence*	460,902	70.13	2.7	2.7
Words, synonyms*	22,311	13.48	2.8	2.8
Protein, S. Cerev*	9,85	1.83	2.5	2.5

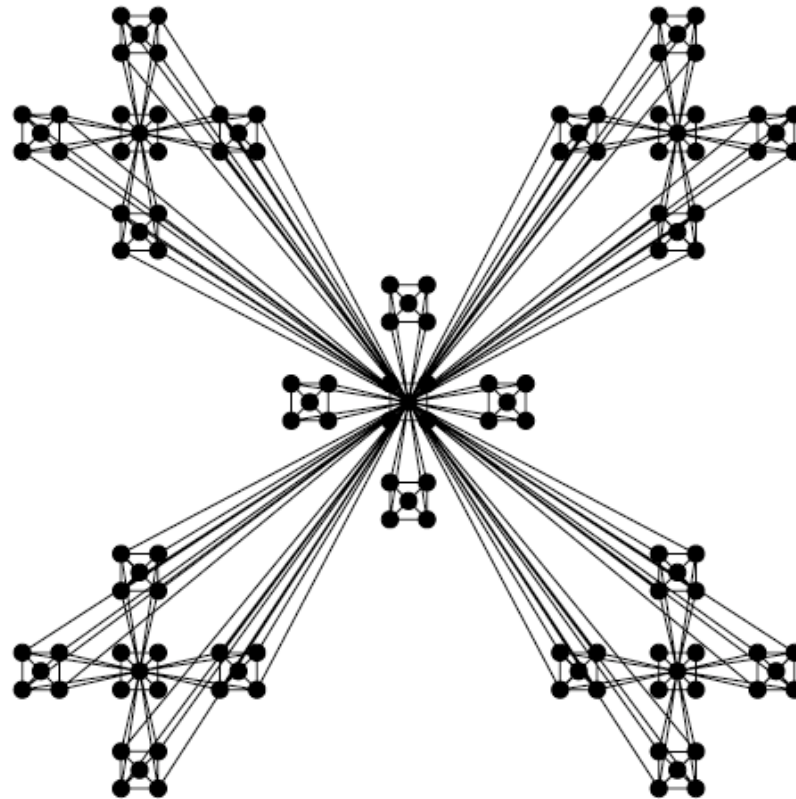
Dinamica di crescita



(a) $n=0, N=5$



(b) $n=1, N=25$



(c) $n=2, N=125$

Cammino medio e clustering

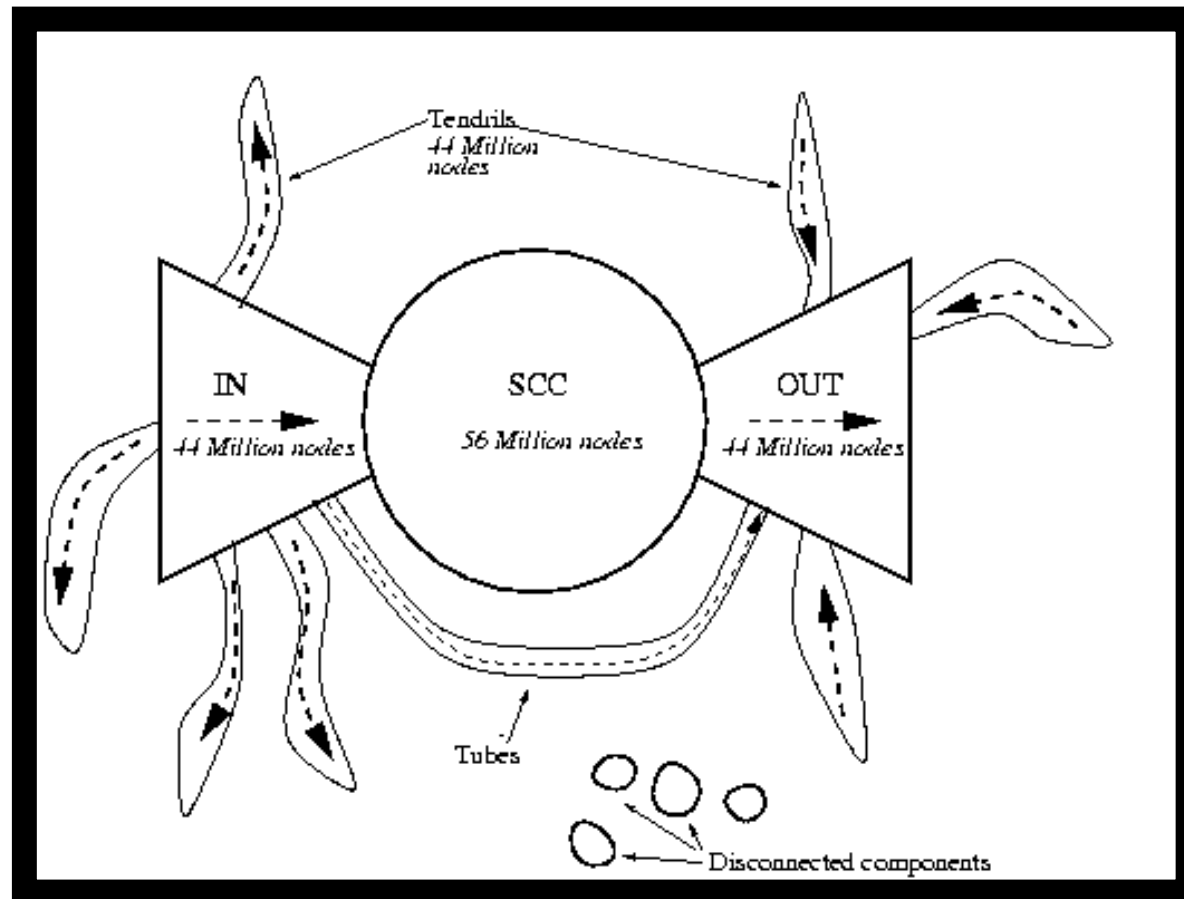
- Cammino medio tra i nodi più breve rispetto a quello di una rete casuale
- Hanno un coefficiente di clustering cinque volte maggiore rispetto a quello dei grafi casuali

I ricchi si arricchiscono?

- Nelle reti reali non solo i nodi più vecchi si arricchiscono di collegamenti
- Per spiegare questo è stato introdotto il “modello a fitness”, Bianconi e Barabasi (2001)
- La fitness è un valore intrinseco del nodo che lo rende capace di competere anche partendo con poche connessioni

$$\Pi_i = f_i k_i / \sum f_i k_j$$

Bowtie



Peculiarità del WWW

- La rete delle risorse accessibili nel WWW ha però una sua peculiarità
 - gli archi non sono simmetrici
 - i link sono dei collegamenti unidirezionali
 - esistono quindi risorse che hanno solo link in entrata e risorse che hanno solo link in uscita
- Per questo i motori di ricerca non riescono a indicizzare l'intero WWW