



ALMA MATER STUDIORUM A.D. 1088

UNIVERSITÀ DI BOLOGNA

AA 2013-2014 - Corso 44756 - 6 CFU

ANALISI DELLE RETI SOCIALI

Social Network Analysis

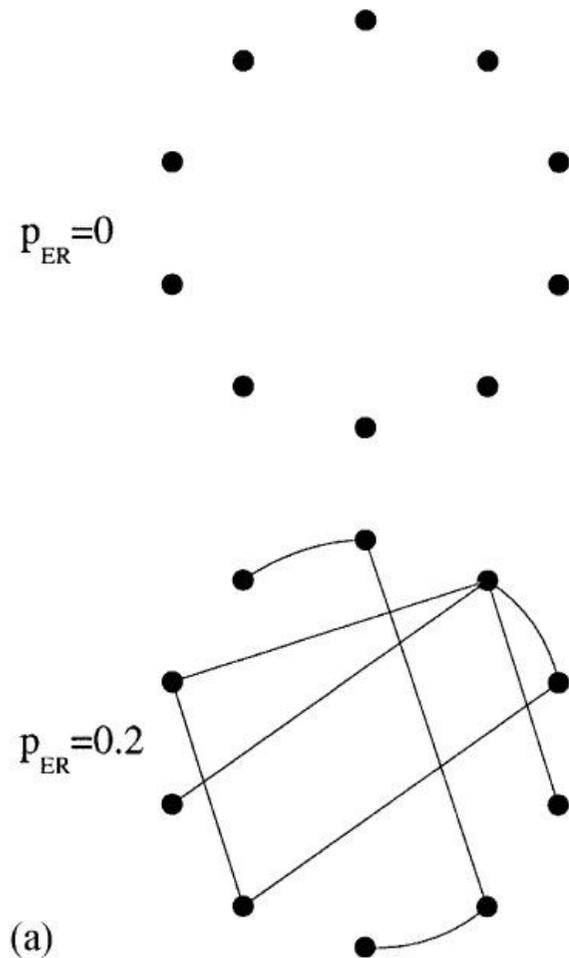
PARTE II – RETI SMALL WORLD E POWER LAW

DISPENSA DI SUPPORTO ALLA DOCENZA

Rev. al 26 settembre 2013

marco ruffino marco.ruffino@unibo.it

RETI DI ERDOS-RENYI



È un modello di generazione random di un network. Un numero di nodi N (10 nella figura) è casualmente unito da n legami, pari a $p_{ER} * [N * (N - 1)/2]$. Nei due esempi, p_{ER} è posta rispettivamente a 0 (nessun legame) ed a 0,2 (9 legami).

Il network random così generato gode di due proprietà:

1. il numero delle sue componenti (ovvero i vari grafi fra loro non connessi, pari a 2 nell'esempio a $p_{ER} = 0,2$) si riduce velocemente appena il *degree* medio dei nodi supera il valore 1;

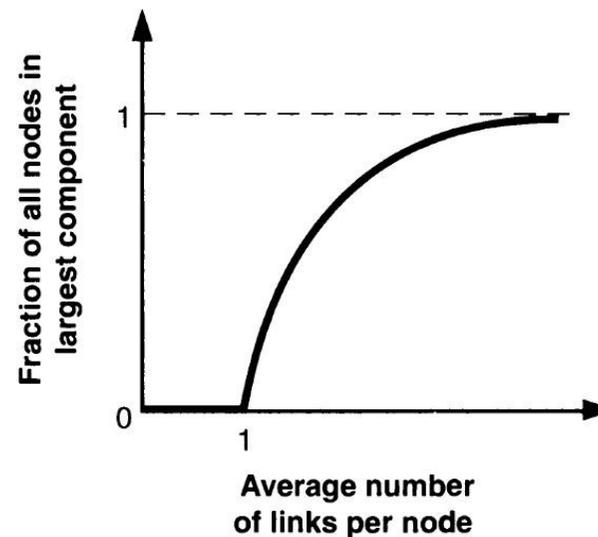


Figure 2.2. Connectivity of a random graph. The fraction of the nodes connected in a single component changes suddenly when the average number of links per node exceeds one.

RETI DI ERDOS-RENYI

2. la probabilità che un nodo abbia un *degree* k è data dalla distribuzione di Poisson¹

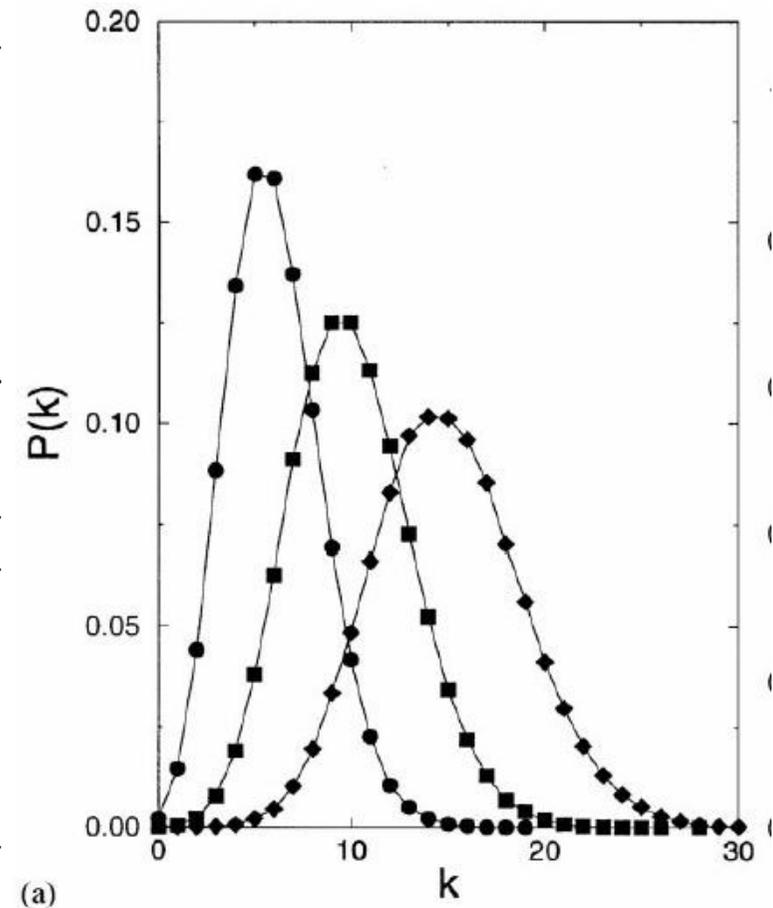
$$P(k) = e^{-\lambda} \lambda^k / k!$$

dove λ è il *degree* medio del network, ovvero $(N-1) * p_{ER}$

Gli esempi in figura, dovuti a Barabàsi (1999), sono relativi (da sinistra verso destra) a $p_{ER}=0,0006$; $p_{ER}=0,001$; $p_{ER}=0,0015$ per $N= 10.000$ (ovvero $\lambda = 6$; 10 ; 15).

Si osservi che le curve sono tracciate solo a fini di miglior comprensione della forma della distribuzione, la stessa essendo discreta (ovvero esistendo solo per k interi, visto che il *degree* di un nodo può essere espresso solo per interi).

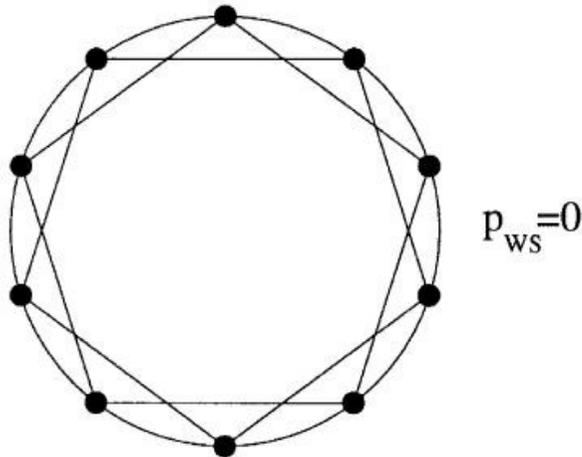
Le reti di Erdos-Renyi non sono un realistico modello dei network reali, in quanto non spiegano la “strana proprietà” della alta raggiungibilità, in presenza di una bassa densità globale.



Fonte: Barabàsi *et al.*, 1999

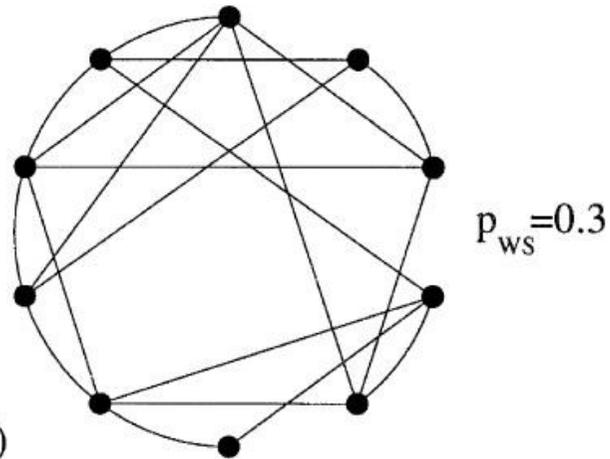
¹ Distribuzione discreta che esprime la probabilità di un insieme di eventi, fra loro indipendenti e che accadono in un determinato intervallo di tempo, essendo noto il loro valor medio. Esempio: λ è il numero medio delle telefonate/ora di un call center; $k= 0, 1, 2, \dots$ è il numero discreto di telefonate per cui calcolare la probabilità dell'evento.

RETI DI WATTS-STROGATZ (o Small Words)



E' un modello alternativo di generazione che cerca di dare una risposta alla empirica rilevazione della elevata raggiungibilità nei network reali.

Si parte da un network ordinato disposto a reticolo come nel grafo a fianco, dove ogni nodo ha *degree* pari a 4. Altrimenti detto, ogni nodo è connesso lungo il cerchio con i suoi due vicini e con i due “vicini dei vicini”.

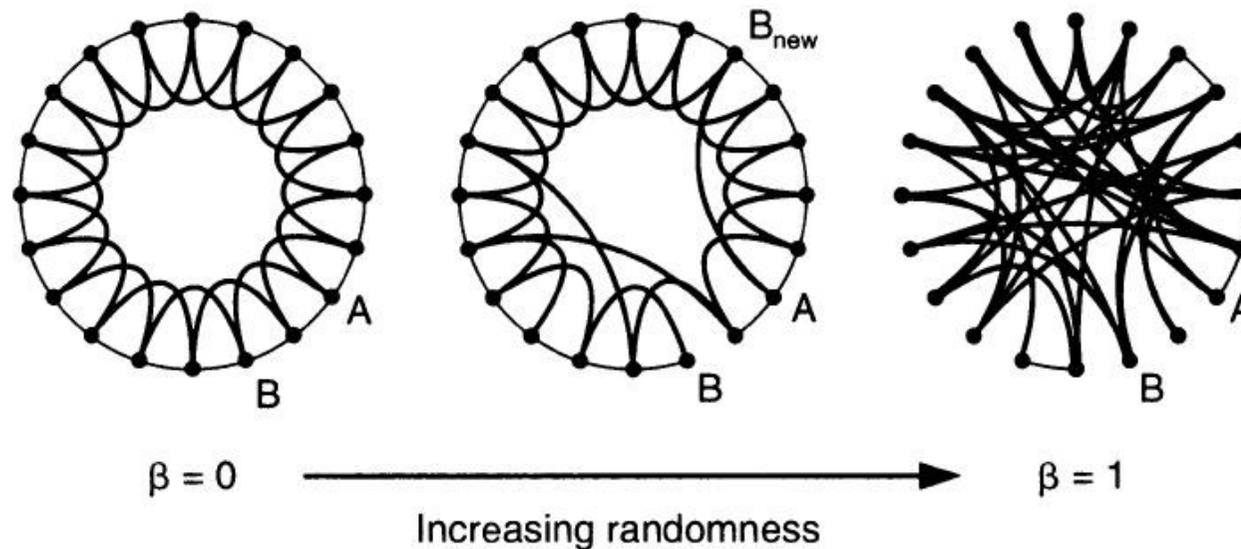


Una frazione p_{ws} dei nodi è successivamente sottoposta a *rewiring* casuale, evitando però che il legame si unisca ad un nodo già direttamente connesso, o al nodo stesso (altrimenti detto, tutto il *rewiring* è produttivo, nel senso che contribuisce sempre a generare una nuova configurazione del network). Nell'esempio (10 nodi e 20 legami), $p_{ws}=0,3$ porta a ricablare casualmente 6 legami.

(b)

RETI DI WATTS-STROGATZ (o Small Words)

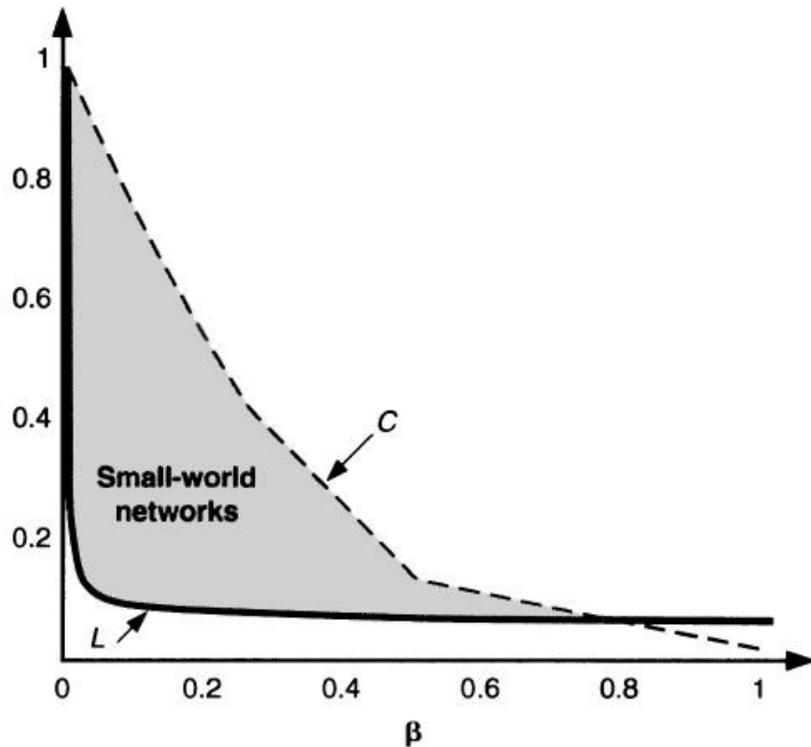
- Per $p_{ws} = \beta = 0$, il reticolo è molto ordinato e la distanza media fra due nodi cresce linearmente per N ;
- Per $p_{ws} = \beta = 1$ il reticolo diviene un grafo random e la distanza media fra due nodi cresce logaritmicamente per N ;
- Nell'intervallo $0 < p_{ws}$ or $\beta < 0,01$ il reticolo assume proprietà *small word*: la distanza media fra due nodi decresce sensibilmente. Fino a p_{ws} or $\beta < 0,5$ si osserva al contempo una lenta discesa del coefficiente di *clustering*.



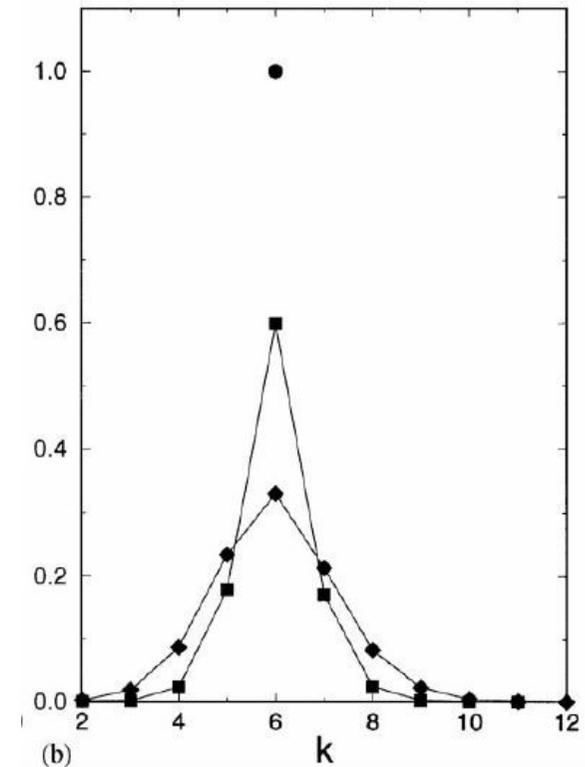
Fonte: Watts, 2003

RETI DI WATTS-STROGATZ (o Small Worlds)

- Come si mostra in figura (dove β va letto equivalente a p_{ws}) l'area grigia ricompresa fra le due curve (distanza media e coefficiente di *clustering*) contiene gli *small worlds*, originariamente caratterizzati dalla nota proprietà dei “sei gradi di separazione” (in Internet interessata da una veloce riduzione, verso il valore 4).
- Anche negli *small worlds*, il *degree* dei nodi segue la distribuzione di Poisson.



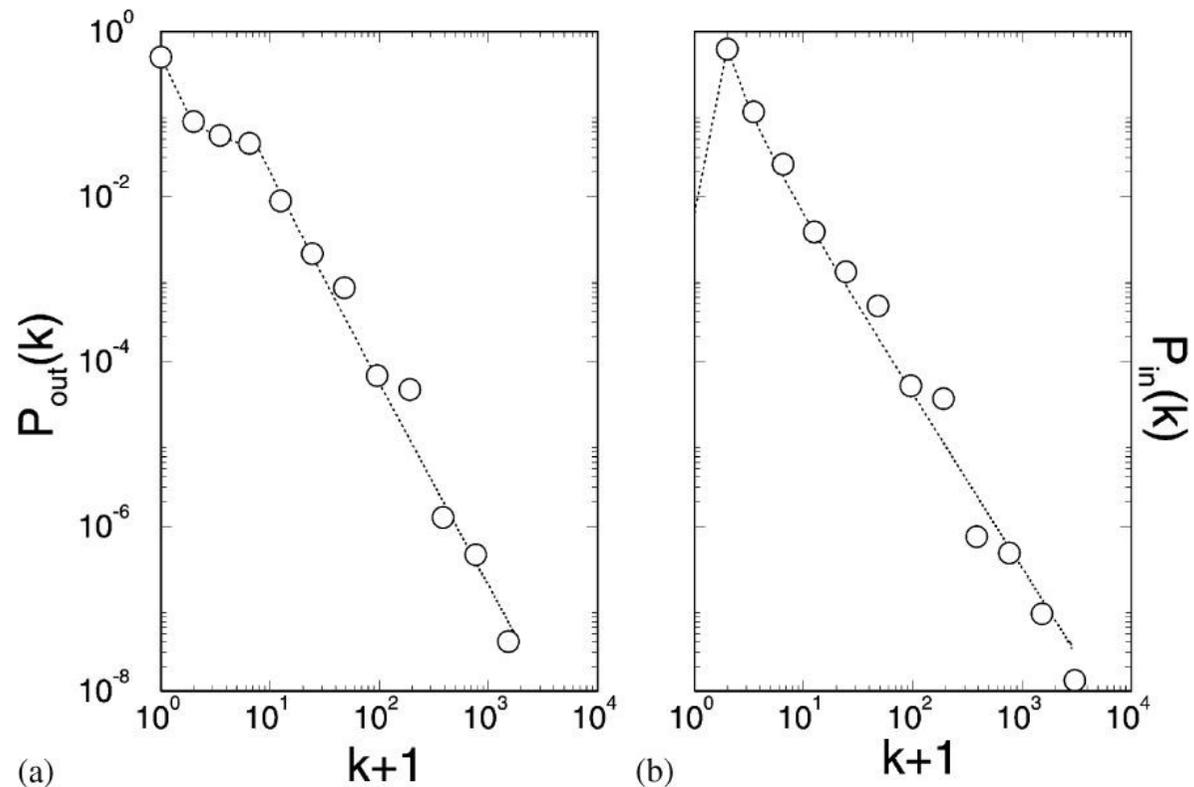
Fonte: Watts, 2003



Fonte: Barabási et al., 1999

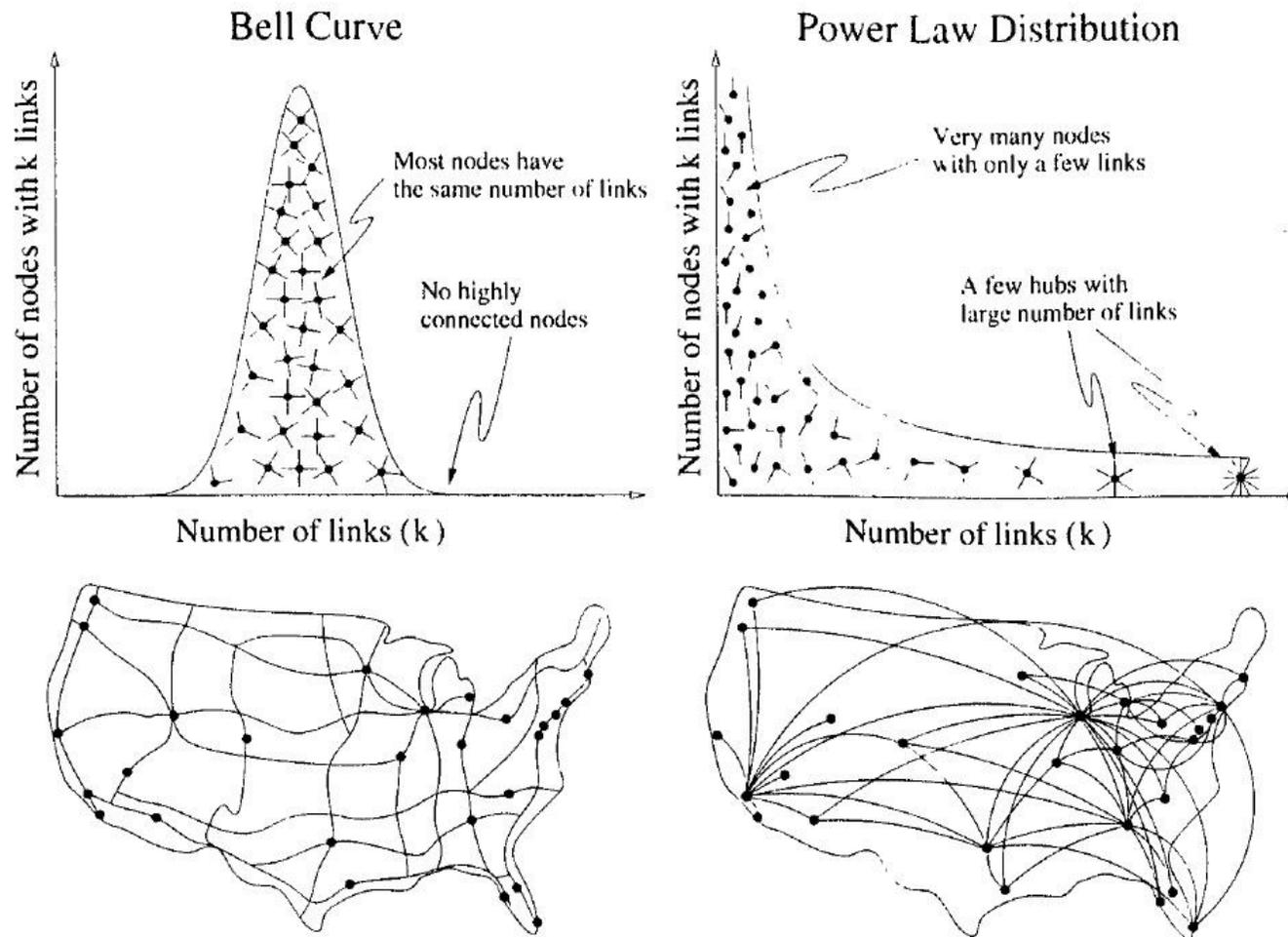
RETI POWER LAW

- In natura , si ritrovano con grande facilità fenomeni che non seguono la distribuzione di Poisson, ma bensì (almeno apparentemente) la legge di potenza $P(k) \sim k^{-\gamma}$
- Nel nostro caso, è come dire che la probabilità $P(k)$ che un nodo in un network sia connesso con k altri nodi (o, se si preferisce, che abbia un *degree* pari a k) decade secondo l'esponente $-\gamma$
- Ciò significa che la distribuzione diviene asimmetrica, con una lunga e pesante coda destra, che contiene pochi nodi con un *degree* maggiore per alcuni ordini di grandezza alla stragrande parte dei nodi restanti.
- Nella figura a fianco (Barabàsi et al. 2000) si mostrano le distribuzioni degli out-e in-degree delle URL trovate analizzando il dominio nd.edu, contenente 325; 729 documenti per 469; 680 links.



RETI POWER LAW

Esempio didattico di rapporto fra topologie di rete e distribuzione del *degree*

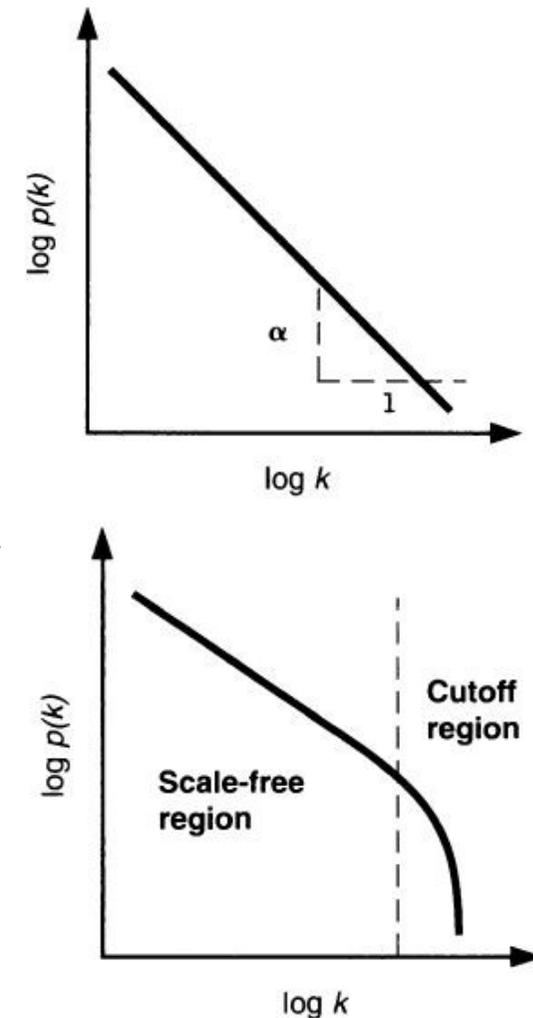


Fonte: Baràbasi (2002)

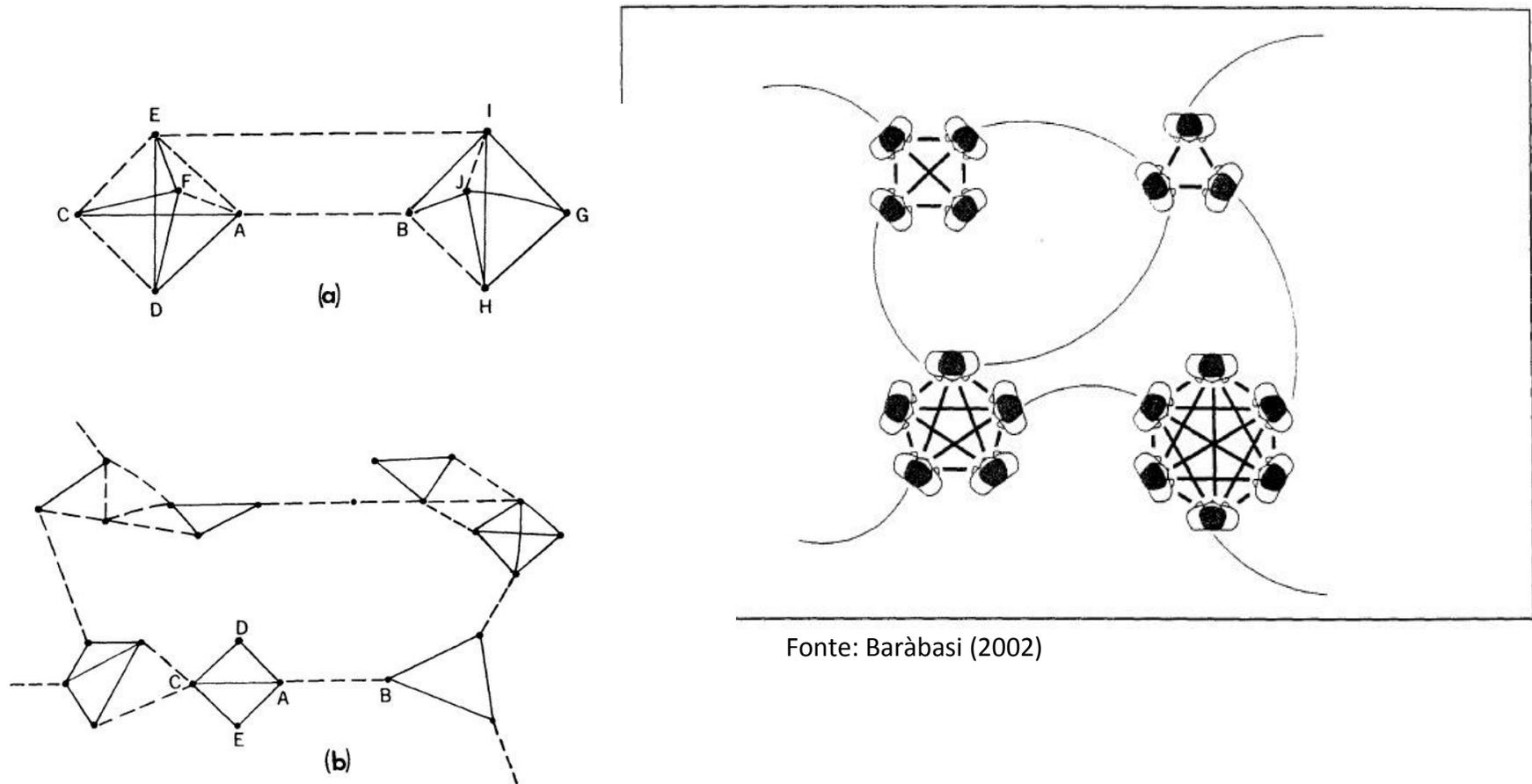
RETI POWER LAW

- Barabàsi e Albert definiscono un algoritmo di generazione della rete basato sulla regola del link preferenziale dei nuovi nodi entrati ai nodi più connessi. In tal modo si generano pochi nodi *hub* di grande dimensione che alimentano direttamente la “*heavy tail*” della distribuzione.
- Una distribuzione *power law* è rappresentata da una retta in un grafico log-log, dove sull’asse x vi è il $\log(k)$ e sull’asse y il logaritmo della probabilità $p(k)$. Le distribuzioni *power law* sono *scale-free*, ovvero mantengono in teoria fino all’infinito la proprietà evidenziata dal grafico log-log.
- La grande “moda” ampiamente verificabile in letteratura sulla diffusione in natura della *power law* si scontra con almeno due problemi, che inducono a riflettere maggiormente:
 - la difficoltà di distinguere in modo certo la distribuzione *power law* da altre (Clauset et al., 2009), fra cui la lognormale, caratterizzate dalla presenza di esponenziali;
 - i limiti “fisici” dei fenomeni umani, che impongono l’esistenza di un valore di *cutoff*, al di sotto del quale la proprietà *scale-free* cessa di presentarsi.

Si tratta dunque di evitare pericolose generalizzazioni.



UNA FAMOSA TEORIA: “ *The Strength of Weak Ties* ”



Fonte: Baràbasi (2002)

FIG. 2.—Local bridges. *a*, Degree 3; *b*, Degree 13. — = strong tie; - - - = weak tie.

Fonte: Granovetter (1973)

BIBLIOGRAFIA ESSENZIALE

ALCUNI TESTI DI RIFERIMENTO

- Barabási A-L., Reka A., Jeong H. (1999), "Mean-field theory for scale-free random networks", *Physica A*, Vol. 272, pp. 173-187
- Barabási A-L., Reka A., Jeong H. (2000), "Scale-free characteristics of random networks: the topology of the world-wide web", *Physica A*, Vol. 281, pp. 69-77.
- Barabási A-L. (2002), *Linked. The new science of networks*, Cambridge (MA), Perseus Publishing.
- Clauset A., Shalizi C. and Newman M. (2009), "Power-law distributions in empirical data", *SIAM review* Vol. 51.4, pp. 661-703.
- Granovetter M. (1973), "The Strength of Weak Ties", *American Journal of Sociology*, Vol. 78, No. 6, pp. 1360-1380.
- Mitzenmacher M., (2004), "A brief history of generative models for power law and lognormal distributions", *Internet mathematics*, Vol. 1, n. 2, pp. 226-251.
- Watts D. (1999), *Small Worlds. The Dynamics of Networks between Order and Randomness*, Princeton, Princeton University Press.
- Watts D. (2003), *Six degrees. The science of a connected age*, New York, Norton & Co.