

Reti complesse

Reti e Misure

Matteo Dell'Amico
dellamico@disi.unige.it

7 novembre 2006



Argomenti Trattati

- 1 Reti complesse
 - Generalità
 - Esempi
 - Obiettivi
- 2 Misurare le reti
 - Distribuzione dei gradi
 - Coefficiente di clustering
 - Fenomeno small-world
- 3 Grafici Casuali
 - Introduzione

Definizioni

Rete

- Collezione di **entità** (nodi) interconnesse.
- Un **arco** (*link*) fra due entità denota un'interazione fra di esse.

Rete complessa

- Rete con topologia **non banale**.

Perché ci interessano le reti?

- Sono **ovunque**.
 - I tradizionali modelli basati su grafici non sono sufficienti.
-
- Le reti su vasta scala **necessitano di nuovi strumenti** per poter essere studiate.
 - “Nuovo” campo scientifico che coinvolge informatica, matematica, fisica, biologia, sociologia ed economia.

Diversi tipi di reti

Le categorie più importanti

- Reti sociali
- Knowledge (Information) Network
- Reti tecnologiche
- Reti biologiche

Reti sociali

Caratteristica

- I link denotano un'interazione sociale

Esempi

- Rete delle conoscenze
- Rete dei co-autori
- Rete delle chiamate telefoniche
- Rete degli attori
- Rete delle E-mail/IM

Knowledge (information) network

Caratteristica

- I nodi memorizzano informazioni, i link le associano

Esempi

- Rete delle citazioni (grafo diretto aciclico)
- World Wide Web (grafo diretto)
- Reti di fiducia (Web of trust)
- Reti P2P

Reti tecnologiche

Caratteristica

- Reti costruite per la distribuzione di “prodotti”

Esempi

- Internet
- Reti per la distribuzione dell'energia
- Reti delle linee aeree
- Reti telefoniche
- Reti di trasporto (strade, ferrovie, ecc...)

Reti biologiche

Caratteristica

- Rappresentano sistemi biologici

Esempi

- Reti d'interazione fra le proteine
- Reti di regolazione dei geni
- Catene alimentari
- Reti neurali

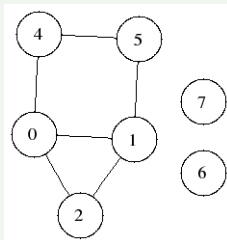
Il mondo è pieno di reti

Che ce ne facciamo?

- Studiamo la loro topologia
- Misuriamo le loro proprietà
- Studiamo la loro evoluzione e la loro dinamica
- Creiamo modelli realistici
- Creiamo algoritmi che sfruttano la struttura delle reti

Il linguaggio dot - grafi indiretti

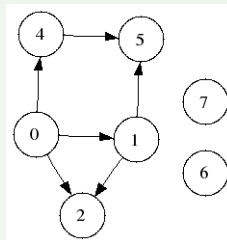
```
graph G {  
  0 -- 1  
  1 -- 2  
  0 -- 2  
  0 -- 4  
  1 -- 5  
  4 -- 5  
  6  
  7  
}
```



- Gli archi sono contrassegnati dall'operatore "--".
- I nodi isolati sono elencati per conto proprio.
- Per ulteriori informazioni: `man dot`, www.graphviz.org.

Il linguaggio dot - grafi diretti

```
digraph G {  
  0 -> 1  
  1 -> 2  
  0 -> 2  
  0 -> 4  
  1 -> 5  
  4 -> 5  
  6  
  7  
}
```



- Gli archi sono contrassegnati dall'operatore "**->**".

Disegnare grafi .dot

neato

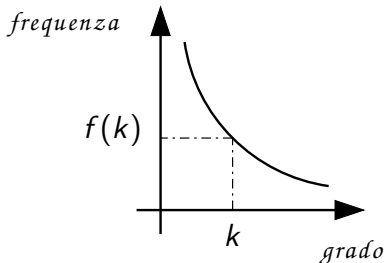
- `neato graph.dot` scrive su `stdout` un file dot annotato con le posizioni dei nodi.
- Le opzioni `-Tformato -ooutfile` permettono di ottenere un'immagine con l'estensione desiderata.
- Le opzioni `-G, -N, -E` impostano attributi di default al grafo, ai nodi ed agli archi rispettivamente.
- Per altre informazioni, ancora `man dot`.

Esempio

- Lo script `draw_dotfile` esegue `"neato -Tsvg -Nstyle=filled -Nlabel= -Nshape=circle -o$2 $1"`



Distribuzione dei gradi

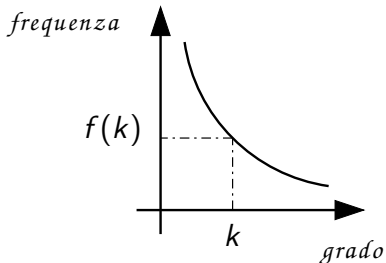


- **Frequenza $f(k)$:** la frazione di nodi con grado k .

Problema

- $f(k)$ è la probabilità che un nodo a caso abbia grado k .
- Determinare la distribuzione di probabilità che meglio approssima i dati osservati.

Distribuzione dei gradi



- **Frequenza $f(k)$:** la frazione di nodi con grado k .

Problema

- $f(k)$ è la probabilità che un nodo a caso abbia grado k .
- Determinare la distribuzione di probabilità che meglio approssima i dati osservati.

Studiare la distribuzione del grado

graph_degree_distribution

- Lo script `graph_degree_distribution` scrive su `stdout`, per ogni linea, un possibile grado e la relativa frequenza.
- Uso: `graph_degree_distribution graph.dot > freq.txt`.

Gnuplot

- Programma interattivo per disegnare grafici.
- Per disegnare: `plot 'datafile' [with lines]`.
- Per creare un file di tipo `eps`: `set term post eps, set output 'nomefile'`.
- Per tornare a disegnare sullo schermo: `set term x11`.
- Scala logaritmica: `setlogscale [x][y], set nologscale`.

Studiare la distribuzione del grado

graph_degree_distribution

- Lo script `graph_degree_distribution` scrive su `stdout`, per ogni linea, un possibile grado e la relativa frequenza.
- Uso: `graph_degree_distribution graph.dot > freq.txt`.

Gnuplot

- Programma interattivo per disegnare grafici.
- Per disegnare: `plot 'datafile' [with lines]`.
- Per creare un file di tipo `eps`: `set term post eps, set output 'nomefile'`.
- Per tornare a disegnare sullo schermo: `set term x11`.
- Scala logaritmica: `set logscale [x][y], set nologscale`.

Power law

- La distribuzione dei gradi in molte reti reali segue la **power law**

$$p(k) \approx c \cdot k^{-\alpha}.$$

Osservazione

- Siamo abituati a ottenere distribuzioni simili alla “distribuzione normale” (gaussiana).
- La power law è **radicalmente diversa**:
 - Ha una frazione non trascurabile di nodi con grado **molto alto** (*hubs*).
 - Non ha “scala caratteristica” (proprietà *scale-free*): il valor medio è **poco informativo**.

Power law

- La distribuzione dei gradi in molte reti reali segue la **power law**

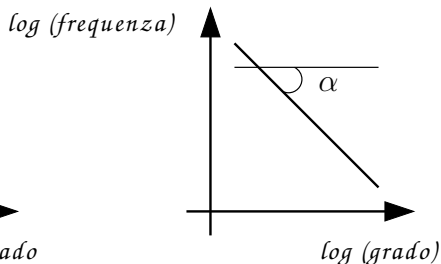
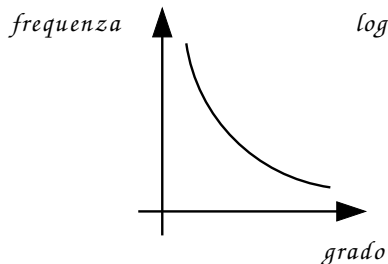
$$p(k) \approx c \cdot k^{-\alpha}.$$

Osservazione

- Siamo abituati a ottenere distribuzioni simili alla “distribuzione normale” (gaussiana).
- La power law è **radicalmente diversa**:
 - Ha una frazione non trascurabile di nodi con grado **molto alto** (*hubs*).
 - Non ha “scala caratteristica” (proprietà *scale-free*): il valor medio è **poco informativo**.

Firma della power law

Firma della power law



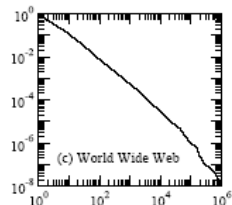
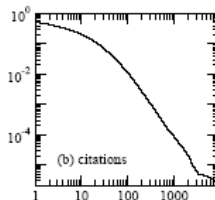
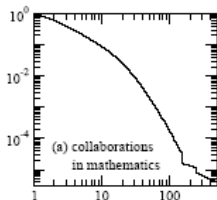
Osservazione

- La distribuzione power law diventa una **retta** se disegnata su scala log-log (bilogaritmica):

$$\log p(k) = -\alpha \log k + \log c.$$

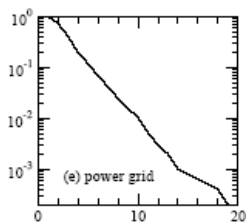
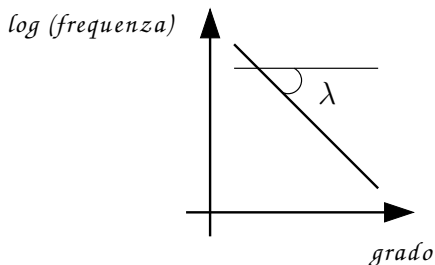
- α è detto **esponente** della power law (spesso, $2 \leq \alpha \leq 3$).

Distribuzione dei gradi in alcuni grafi reali



- Per verificare se una rete è power-law, è meglio valutare la **densità cumulativa** della distribuzione dei grafi.
- Dato $p(k)$, calcoliamo $P(k) = \sum_{i \geq k} p(i)$.
- Otteniamo sempre una retta, con coefficiente angolare $-\alpha + 1$.
- Lo script `cumulative_degree_distribution` genera la distribuzione di densità cumulativa a partire dall'output di `graph_degree_distribution`.

Distribuzione esponenziale



Definizione

- $p(k) \approx c \cdot e^{-\lambda k}$
- Identificata da una retta se disegnata su scala log-lineare

$$\log p(k) = -\lambda k + \log c.$$

Statistiche collettive

	network	type	n	m	z	ℓ	α	$C^{(1)}$	$C^{(2)}$	r	Ref(s).
social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208	20, 416
	company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88	0.276	105, 323
	math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34	0.120	107, 182
	physics coauthorship	undirected	52 909	245 900	9.27	6.19	–	0.45	0.56	0.363	311, 313
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60	0.127	311, 313
	telephone call graph	undirected	47 000 000	80 000 000	3.16	–	2.1	–	–	–	8, 9
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0	–	0.16	–	136
	email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13	0.092	321
	student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001	–0.029	45
	sexual contacts	undirected	2 810	–	–	–	3.2	–	–	–	265, 266
information	WWW nd.edu	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	–0.067	14, 34
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7	–	–	–	74
	citation network	directed	783 339	6 716 198	8.57	–	3.0/–	–	–	–	351
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	–	0.13	0.15	0.157	244
	word co-occurrence	undirected	460 902	17 000 000	70.13	–	2.7	–	0.44	–	119, 157
technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39	–0.189	86, 148
	power grid	undirected	4941	6 504	2.67	18.99	–	0.10	0.080	–0.003	416
	train routes	undirected	587	19 603	66.79	2.16	–	–	0.69	–0.033	366
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082	–0.016	318
	software classes	directed	1 377	2 213	1.61	1.51	–	0.033	0.012	–0.119	395
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030	–0.154	155
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011	–0.366	6, 354
	metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.090	0.67	–0.240	214
biological	protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.072	0.071	–0.156	212
	marine food web	directed	135	598	4.43	2.05	–	0.16	0.23	–0.263	204
	freshwater food web	directed	92	997	10.84	1.90	–	0.20	0.087	–0.326	272
	neural network	directed	307	2 359	7.68	3.97	–	0.18	0.28	–0.226	416, 421

TABLE II Basic statistics for a number of published networks. The properties measured are: type of graph, directed or undirected; total number of vertices n ; total number of edges m ; mean degree z ; mean vertex-vertex distance ℓ ; exponent α of degree distribution if the distribution follows a power law (or $^{\text{in/out}}$ if not; in/out-degree exponents are given for directed graphs); clustering coefficient $C^{(1)}$ from Eq. (3); clustering coefficient $C^{(2)}$ from Eq. (6); and degree correlation coefficient r , Sec. III.F. The last column gives the citation(s) for the network in the bibliography. Blank entries indicate unavailable data.

Clustering

- Molte reti hanno la proprietà di essere composte da gruppi di nodi molto ben connessi.
- Nelle reti sociali, il **clustering (transitività)** rappresenta il principio secondo il quale due miei amici hanno una buona probabilità di essere amici tra di loro.
- Reti con alto clustering hanno un elevato numero di **triangoli**.

Coefficienti di clustering

Definizione

$$C^{(1)}(G) = \frac{\text{numero di triangoli in } G}{\text{triple connesse in } G}$$

Definizione

$$c_i = \frac{\text{triangoli centrati nel nodo } i}{\text{triple centrate nel nodo } i}$$

$$C^{(2)}(G) = \frac{\sum_i c_i}{n}$$

Coefficienti di clustering

Definizione

$$C^{(1)}(G) = \frac{\text{numero di triangoli in } G}{\text{triple connesse in } G}$$

Definizione

$$c_i = \frac{\text{triangoli centrati nel nodo } i}{\text{triple centrate nel nodo } i}$$

$$C^{(2)}(G) = \frac{\sum_i c_i}{n}$$

Coefficienti di clustering

Definizione

$$C^{(1)}(G) = \frac{\text{numero di triangoli in } G}{\text{triple connesse in } G}$$

Definizione

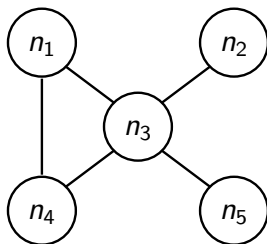
$$c_i = \frac{\text{triangoli centrati nel nodo } i}{\text{triple centrate nel nodo } i}$$

$$C^{(2)}(G) = \frac{\sum_i c_i}{n}$$

Osservazione

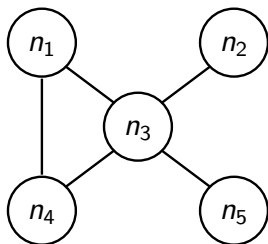
- $C^{(1)}$ assegna più peso ai nodi di grado alto, $C^{(2)}$ assegna lo stesso peso a tutti.

Esempio



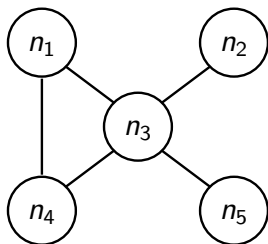
- `clustering dotfile` stampa i valori di $C^{(1)}$ e $C^{(2)}$.

Esempio



- clustering *dotfile* stampa i valori di $C^{(1)}$ e $C^{(2)}$.
- $C^{(1)} = \frac{3}{1+1+6} = \frac{3}{8} \approx 0.37$.

Esempio



- clustering *dotfile* stampa i valori di $C^{(1)}$ e $C^{(2)}$.
- $C^{(1)} = \frac{3}{1+1+6} = \frac{3}{8} \approx 0.37$.
- $C^{(2)} = \frac{1}{5} \left(1 + 1 + \frac{1}{6}\right) = \frac{13}{30} \approx 0.43$.

Statistiche collettive

	network	type	n	m	z	ℓ	α	$C^{(1)}$	$C^{(2)}$	r	Ref(s).
social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208	20, 416
	company directors	undirected	7 673	55 392	14.44	4.60	-	0.59	0.88	0.276	105, 323
	math coauthorship	undirected	253 339	496 489	3.92	7.57	-	0.15	0.34	0.120	107, 182
	physics coauthorship	undirected	52 909	245 300	9.27	6.19	-	0.45	0.56	0.363	311, 313
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	-	0.088	0.60	0.127	311, 313
	telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1				8, 9
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16		136
	email address books	directed	16 881	57 029	3.38	5.22	-	0.17	0.13	0.092	321
	student relationships	undirected	573	477	1.66	16.01	-	0.005	0.001	-0.029	45
	sexual contacts	undirected	2 810				3.2				265, 266
information	WWW nd.edu	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	-0.067	14, 34
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7				74
	citation network	directed	783 339	6 716 198	8.57		3.0/-				351
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	-	0.13	0.15	0.157	244
	word co-occurrence	undirected	460 902	17 000 000	70.13		2.7		0.44		119, 157
technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39	-0.189	86, 148
	power grid	undirected	4941	6 504	2.67	18.99	-	0.10	0.080	-0.003	416
	train routes	undirected	587	19 603	66.79	2.16	-	-	0.69	-0.033	366
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082	-0.016	318
	software classes	directed	1 377	2 213	1.61	1.51	-	0.033	0.012	-0.119	395
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030	-0.154	155
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011	-0.366	6, 354
biological	metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.090	0.67	-0.240	214
	protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.072	0.071	-0.156	212
	marine food web	directed	135	598	4.43	2.05	-	0.16	0.23	-0.263	204
	freshwater food web	directed	92	997	10.84	1.90	-	0.20	0.087	-0.326	272
	neural network	directed	307	2 359	7.68	3.97	-	0.18	0.28	-0.226	416, 421

TABLE II Basic statistics for a number of published networks. The properties measured are: type of graph, directed or undirected; total number of vertices n ; total number of edges m ; mean degree z ; mean vertex-vertex distance ℓ ; exponent α of degree distribution if the distribution follows a power law (or 1.7 if not; in/out-degree exponents are given for directed graphs); clustering coefficient $C^{(1)}$ from Eq. (3); clustering coefficient $C^{(2)}$ from Eq. (6); and degree correlation coefficient r , Sec. III.F. The last column gives the citation(s) for the network in the bibliography. Blank entries indicate unavailable data.

Sei gradi di separazione



Esperimento di Milgram (1967)

Determinare la lunghezza media delle catene di conoscenti che collegano due persone (che non si conoscono) negli USA

Tipica istanza del problema

- Ad una persona “sorgente” (in Nebraska) viene data una lettera da **consegnare** ad una persona “destinazione” (a Boston).
- La sorgente ha informazioni di base sulla destinazione (indirizzo, occupazione, ecc...).

Sei gradi di separazione



Esperimento di Milgram (1967)

Determinare la lunghezza media delle catene di conoscenti che collegano due persone (che non si conoscono) negli USA

Tipica istanza del problema

- Ad una persona “sorgente” (in Nebraska) viene data una lettera da **consegnare** ad una persona “destinazione” (a Boston).
- La sorgente ha informazioni di base sulla destinazione (indirizzo, occupazione, ecc. . .).

Esperimento di Milgram

Dettagli

- La sorgente deve inviare la lettera ad un **conoscente diretto** (ricevente) provando a raggiungere la destinazione nel minor numero di passi possibile.
- I riceventi seguono la stessa regola della sorgente.
- La lettera passa di ricevente in ricevente finché non arriva a destinazione o non si *perde* (!)

Risultato

- La lunghezza media di una catena di conoscenze *che raggiunge la destinazione* è circa **sei**.

Esperimento di Milgram

Dettagli

- La sorgente deve inviare la lettera ad un **conoscente diretto** (ricevente) provando a raggiungere la destinazione nel minor numero di passi possibile.
- I riceventi seguono la stessa regola della sorgente.
- La lettera passa di ricevente in ricevente finché non arriva a destinazione o non si *perde* (!)

Risultato

- La lunghezza media di una catena di conoscenze *che raggiunge la destinazione* è circa **sei**.

Esperimento di Milgram

Dettagli

- La sorgente deve inviare la lettera ad un **conoscente diretto** (ricevente) provando a raggiungere la destinazione nel minor numero di passi possibile.
- I riceventi seguono la stessa regola della sorgente.
- La lettera passa di ricevente in ricevente finché non arriva a destinazione o non si *perde* (!)

Risultato

- La lunghezza media di una catena di conoscenze *che raggiunge la destinazione* è circa **sei**.

Esperimento di Milgram

Dettagli

- La sorgente deve inviare la lettera ad un **conoscente diretto** (ricevente) provando a raggiungere la destinazione nel minor numero di passi possibile.
- I riceventi seguono la stessa regola della sorgente.
- La lettera passa di ricevente in ricevente finché non arriva a destinazione o non si *perde* (!)

Risultato

- La lunghezza media di una catena di conoscenze *che raggiunge la destinazione* è circa **sei**.

Misurare il fenomeno small-world

Distanza geodesica

d_{ij} = lunghezza del cammino minimo fra i nodi i e j .

Diametro

$d = \max_{i,j} \{d_{ij} : i, j \text{ connessi}\}$.

Media della distanza geodesica

$$\ell = \frac{1}{n(n-1)/2} \sum_{i>j} d_{ij}$$

Problema

Come trattare le coppie di nodi disconnesse?

Misurare il fenomeno small-world

Distanza geodesica

d_{ij} = lunghezza del cammino minimo fra i nodi i e j .

Diametro

$d = \max_{i,j} \{d_{ij} : i, j \text{ connessi}\}$.

Media della distanza geodesica

$$\ell = \frac{1}{n(n-1)/2} \sum_{i>j} d_{ij}$$

Problema

Come trattare le coppie di nodi disconnesse?

Misurare il fenomeno small-world

Distanza geodesica

d_{ij} = lunghezza del cammino minimo fra i nodi i e j .

Diametro

$d = \max_{i,j} \{d_{ij} : i, j \text{ connessi}\}$.

Media della distanza geodesica

$$\ell = \frac{1}{n(n-1)/2} \sum_{i>j} d_{ij}$$

Problema

Come trattare le coppie di nodi disconnesse?

Misurare il fenomeno small-world

Distanza geodesica

d_{ij} = lunghezza del cammino minimo fra i nodi i e j .

Diametro

$d = \max_{i,j} \{d_{ij} : i, j \text{ connessi}\}$.

Media della distanza geodesica

$$\ell = \frac{1}{n(n-1)/2} \sum_{i>j} d_{ij}$$

Media armonica della distanza geodesica

$$\ell^{-1} = \frac{1}{n(n-1)/2} \sum_{i>j} d_{ij}^{-1}$$

Misurare il fenomeno small-world

- `distances dotfile` stampa diametro, distanza media e media armonica delle distanze.

Statistiche collettive

	network	type	n	m	z	ℓ	α	$C^{(1)}$	$C^{(2)}$	r	Ref(s).
social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208	20, 416
	company directors	undirected	7 673	55 392	14.44	4.60	-	0.59	0.88	0.276	105, 323
	math coauthorship	undirected	253 339	496 489	3.92	7.57	-	0.15	0.34	0.120	107, 182
	physics coauthorship	undirected	52 909	245 300	9.27	6.19	-	0.45	0.56	0.363	311, 313
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	-	0.088	0.60	0.127	311, 313
	telephone call graph	undirected	47 000 000	80 000 000	3.16	-	2.1	-	-	-	8, 9
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0	-	0.16	-	136
	email address books	directed	16 881	57 029	3.38	5.22	-	0.17	0.13	0.092	321
	student relationships	undirected	573	477	1.66	16.01	-	0.005	0.001	-0.029	45
	sexual contacts	undirected	2 810	-	-	-	3.2	-	-	-	265, 266
information	WWW nd.edu	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	-0.067	14, 34
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7	-	-	-	74
	citation network	directed	783 339	6 716 198	8.57	-	3.0/-	-	-	-	351
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	-	0.13	0.15	0.157	244
	word co-occurrence	undirected	460 902	17 000 000	70.13	-	2.7	-	0.44	-	119, 157
technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39	-0.189	86, 148
	power grid	undirected	4941	6 504	2.67	18.99	-	0.10	0.080	-0.003	416
	train routes	undirected	587	19 603	66.79	2.16	-	-	0.69	-0.033	366
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082	-0.016	318
	software classes	directed	1 377	2 213	1.61	1.51	-	0.033	0.012	-0.119	395
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030	-0.154	155
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011	-0.366	6, 354
	metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.090	0.67	-0.240	214
biological	protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.072	0.071	-0.156	212
	marine food web	directed	135	598	4.43	2.05	-	0.16	0.23	-0.263	204
	freshwater food web	directed	92	997	10.84	1.90	-	0.20	0.087	-0.326	272
	neural network	directed	307	2 359	7.68	3.97	-	0.18	0.28	-0.226	416, 421

TABLE II Basic statistics for a number of published networks. The properties measured are: type of graph, directed or undirected; total number of vertices n ; total number of edges m ; mean degree z ; mean vertex-vertex distance ℓ ; exponent α of degree distribution if the distribution follows a power law (or $^{\infty}$ if not; in/out-degree exponents are given for directed graphs); clustering coefficient $C^{(1)}$ from Eq. (3); clustering coefficient $C^{(2)}$ from Eq. (6); and degree correlation coefficient r , Sec. III.F. The last column gives the citation(s) for the network in the bibliography. Blank entries indicate unavailable data.

Una distanza geodesica corta è sufficiente?

Osservazione

- Deve esistere un metodo per sfruttare i percorsi brevi in condizioni “realistiche”.
- I link dei vicini devono essere, in qualche modo, prevedibili: un amico avvocato può conoscere molti altri avvocati. **Il clustering alza il diametro ma aiuta la navigabilità.**
- L'esperimento di Milgram dimostra che **le persone sono capaci a “navigare” sulla rete sociale.**

Risposta

- Si parla di fenomeno small-world quando si presentano **basso diametro, facilità di navigazione ed elevato coefficiente di clustering.**

Una distanza geodesica corta è sufficiente?

Osservazione

- Deve esistere un metodo per sfruttare i percorsi brevi in condizioni “realistiche”.
- I link dei vicini devono essere, in qualche modo, prevedibili: un amico avvocato può conoscere molti altri avvocati. **Il clustering alza il diametro ma aiuta la navigabilità.**
- L'esperimento di Milgram dimostra che **le persone sono capaci a “navigare” sulla rete sociale.**

Risposta

- Si parla di fenomeno small-world quando si presentano **basso diametro, facilità di navigazione ed elevato coefficiente di clustering.**

Una distanza geodesica corta è sufficiente?

Osservazione

- Deve esistere un metodo per sfruttare i percorsi brevi in condizioni “realistiche”.
- I link dei vicini devono essere, in qualche modo, prevedibili: un amico avvocato può conoscere molti altri avvocati. **Il clustering alza il diametro ma aiuta la navigabilità.**
- L’esperimento di Milgram dimostra che **le persone sono capaci a “navigare” sulla rete sociale.**

Risposta

- Si parla di fenomeno small-world quando si presentano **basso diametro, facilità di navigazione ed elevato coefficiente di clustering.**

Una distanza geodesica corta è sufficiente?

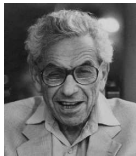
Osservazione

- Deve esistere un metodo per sfruttare i percorsi brevi in condizioni “realistiche”.
- I link dei vicini devono essere, in qualche modo, prevedibili: un amico avvocato può conoscere molti altri avvocati. **Il clustering alza il diametro ma aiuta la navigabilità.**
- L’esperimento di Milgram dimostra che **le persone sono capaci a “navigare” sulla rete sociale.**

Risposta

- Si parla di fenomeno small-world quando si presentano **basso diametro, facilità di navigazione ed elevato coefficiente di clustering.**

I grafici casuali di Erdős–Rényi



Il modello $G_{n,p}$

- n : numero di vertici
- $0 \leq p \leq 1$
- Per ogni coppia di nodi i e j , genera l'arco (i, j) con probabilità p in modo **indipendente**.

Componenti connesse

- `erdos_renyi_graph n p dotfile` genera un grafo casuale $G_{n,p}$ in dotfile.
- Lo script `color_graph_components dot1 dot2` colora le componenti connesse di colori differenti.

Per il 7 novembre

- Creare grafici casuali e studiarne le caratteristiche:
 - distribuzione del grado
 - coefficiente di clustering
 - diametro e distanze medie
 - componenti connesse
- Scrivere una breve relazione e sottometerla su aulaweb.

Domande

- La distribuzione del grado è una power law?
- Che valori assume il coefficiente di clustering rispetto alle reti reali viste negli esempi?
- Il diametro ci permetterebbe di parlare di rete “small world”?
- Per quali valori di n e p si ottiene una componente connessa di dimensioni paragonabili all'intera rete (*componente gigante*)?
- Per quali valori di n e p tutti i nodi fanno parte di un'unica componente connessa?
- Come si evolvono le caratteristiche delle reti fissando p al crescere di $n \rightarrow \infty$?
- Come si evolvono le caratteristiche delle reti al crescere di $n \rightarrow \infty$ fissando il grado medio np ?