

Symulacje sieci w Zespole Układów Złożonych

Seminarium Wydziału Fizyki i Informatyki Stosowanej AGH, 9 marca 2012

plan

- Sieć Erdösa-Rényi,
- Macierz połączeń
- Sieć Watts-Strogatza
- Współczynnik klasteryzacji, odległość, średnica
- Sieci rosnące
- Struktura, asortatywność
- Model konfiguracyjny
- Grafy krawędziowe
- Sieci w ZUZ
- Przykład danych: LiveJournal jako graf krawędziowy

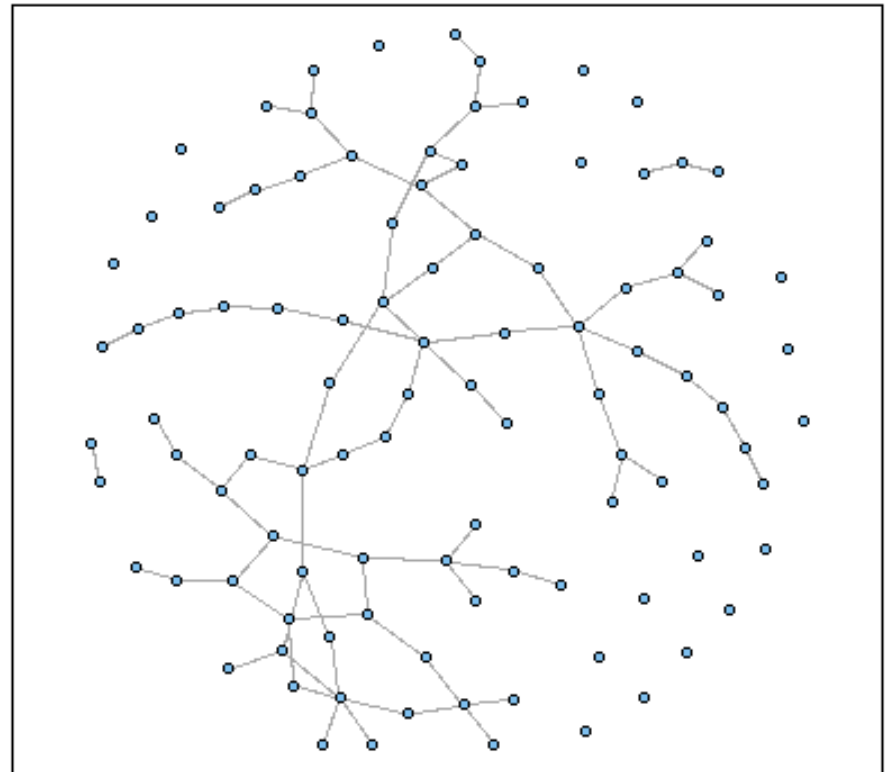
Sieci Erdösa-Rényi

- Weź N wierzchołków
- Każdą parę wierzchołków połącz wiązaniem z p -stwem p

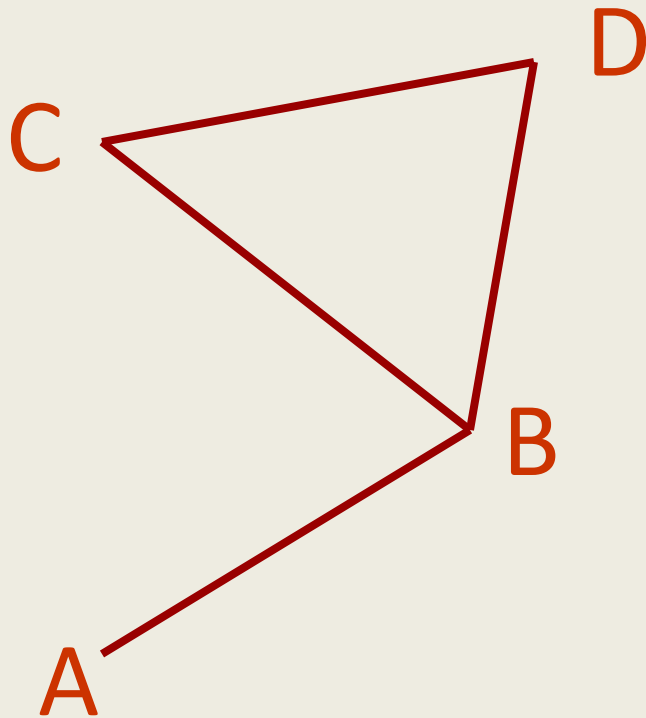
Średni stopień wierzchołka
 $\langle k \rangle = (N-1)p$

Rozkład stopnia wierzchołka

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$



Macierz połączeń C_{ik}



0	1	0	0
1	0	1	1
0	1	0	1
0	1	1	0

Inne opcje: wagi, graf skierowany
Wada: zliczanie macierzy

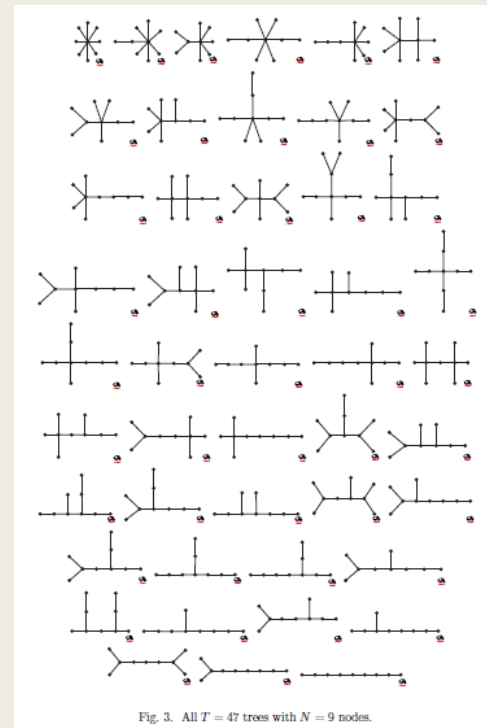


Fig. 3. All $T = 47$ trees with $N = 9$ nodes.

[dla drzew $N=17$, $\#=48629$;
 S. Piec e a, IJMPC 2005]

Współczynnik klasteryzacji C

Niech wierzchołek i o stopniu k_i ma L_i wiązań między swoimi sąsiadami.

$$C = \frac{1}{N} \sum_{i=1}^N \frac{2L_i}{k_i(k_i - 1)}$$

Średnica grafu

„Najdłuższa z najkrótszych ścieżek”, gdzie najkrótsza ścieżka $d(i, j)$ = najmniejsza ilość wiązań między wierzchołkami i, j

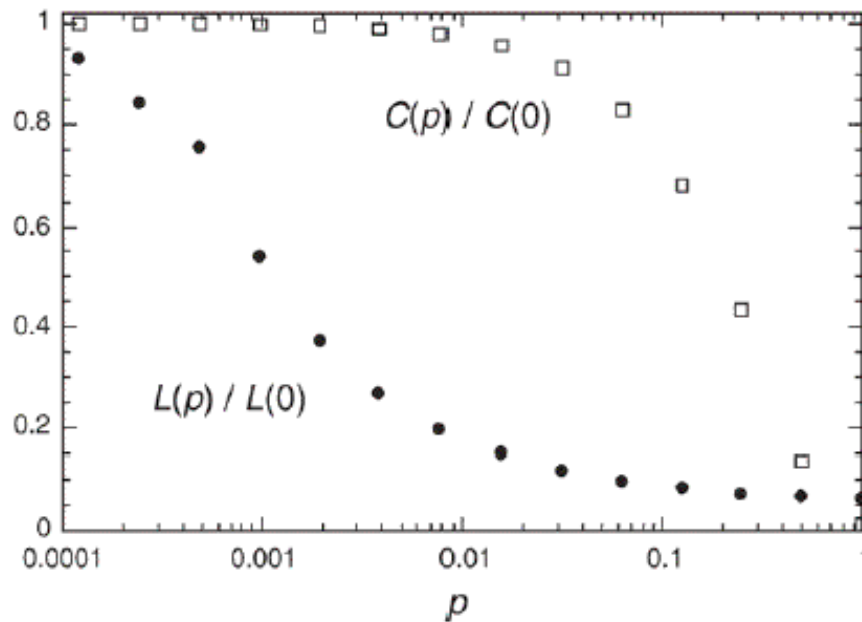
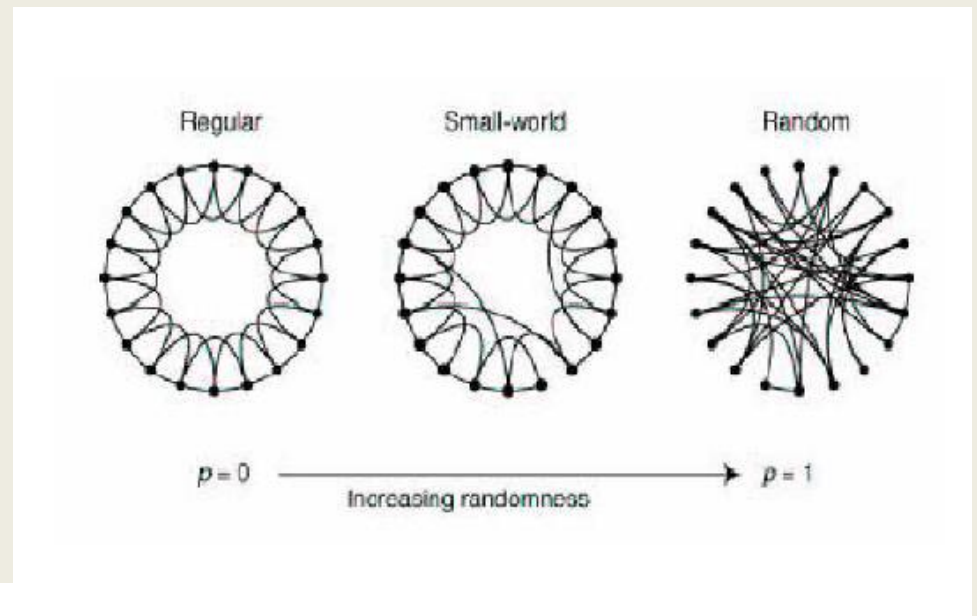
„Średnia z najkrótszych ścieżek”

$$d = \frac{1}{N(N-1)} \sum_{i, j=1}^N d(i, j)$$

Własność „małego świata” $d(N) \propto \ln(N)$

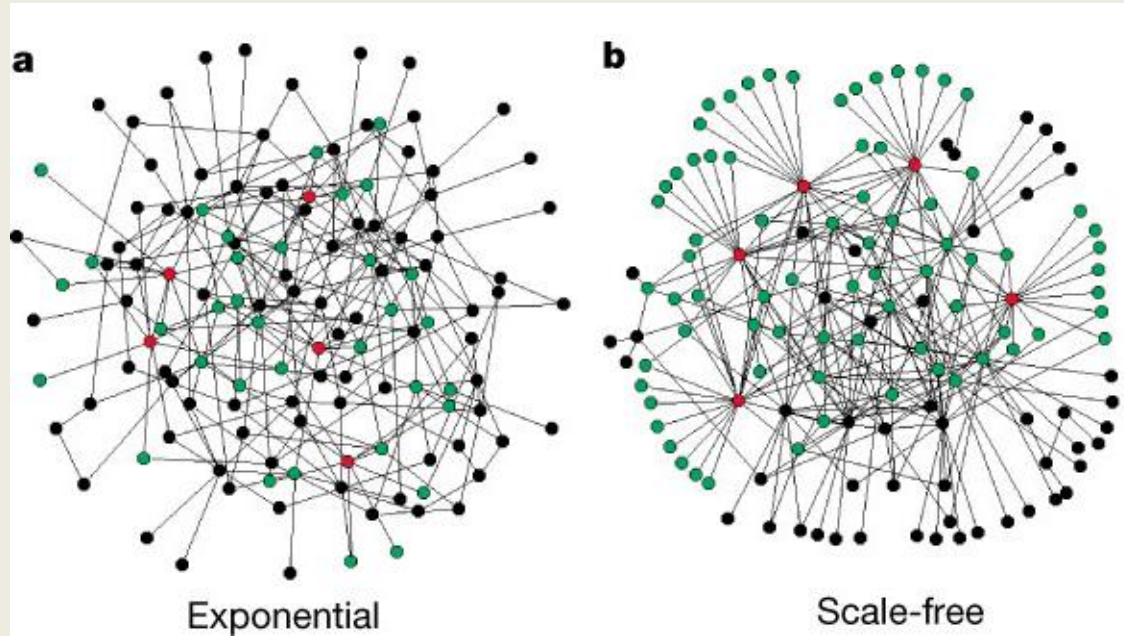
bo $N \approx \langle k \rangle^{d(N)}$

Sieć Wattsa- Strogatza



Sieci rosnące

- Weź M wierzchołków połączonych ze sobą
- dołączaj nowe wierzchołki, każdy za pomocą M wiązań



eksponencjalne

bezskałowe*

$$P(k) \propto \left(\frac{M}{M+1} \right)^k$$

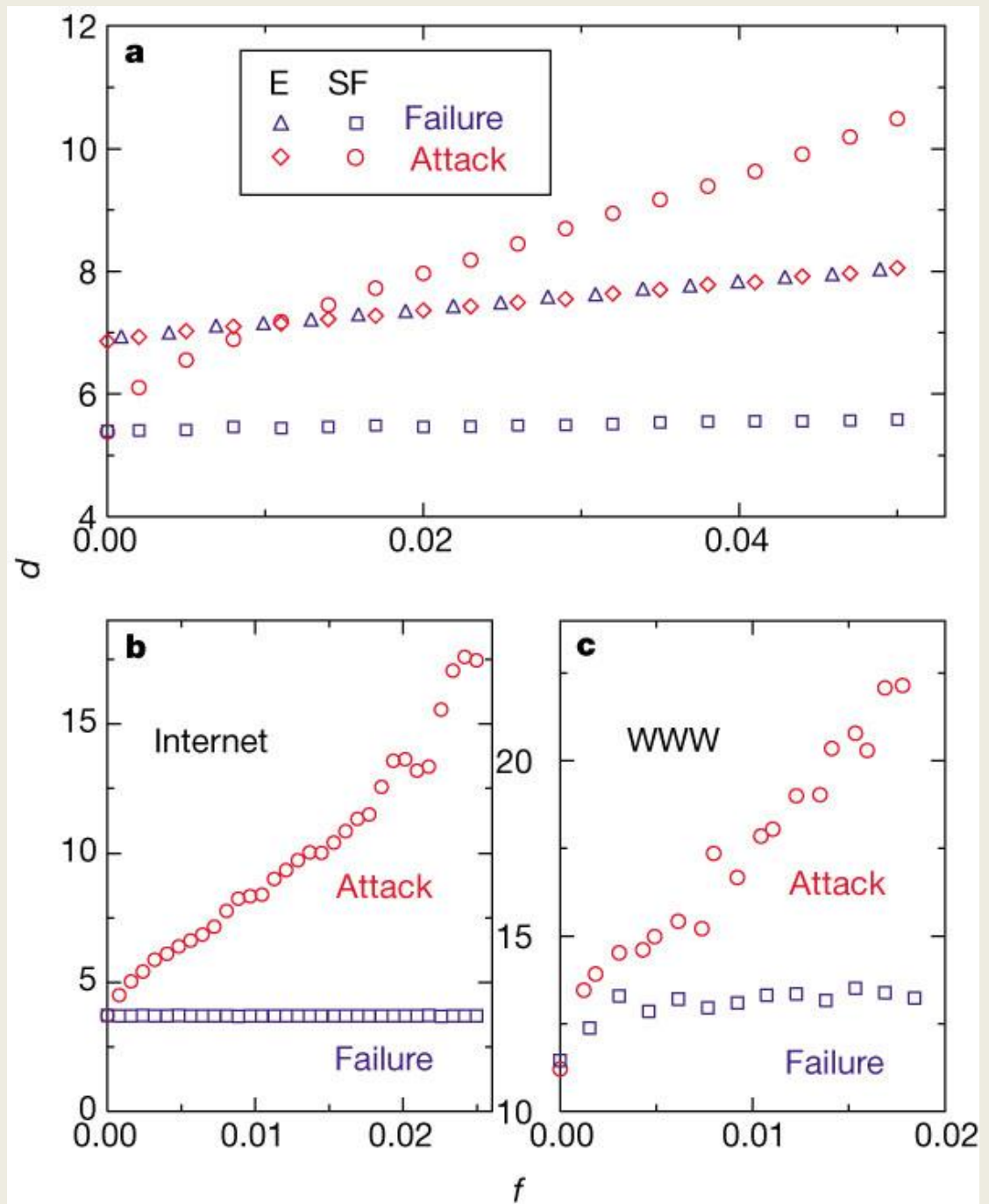
$$P(k) \propto k^{-\gamma}$$

*Dołączanie z preferencją = dołączanie do wierzchołków z p -stwem proporcjonalnym do ich stopnia

SIEĆ	N	#	d	C	γ
aktorzy	45×10^4	25×10^6	3.48	0.78	2.3
www Altavista	2×10^9	2×10^{10}	16.18	-	2.1/2.7
Współautorzy – mat *	25×10^4	50×10^4	7.57	0.34	-
Współautorzy - fiz	5×10^4	25×10^4	6.19	0.56	-
Rozmowy telefoniczne	47×10^6	8×10^7	-	-	2.1
Łańcuchy pokarmowe w wodzie	92	997	1.9	0.087	-
Oddziaływania białek	2115	2240	6.8	0.071	2.4
Kontakty seksualne	2810	-	-	-	3.2
Słowa w zdaniach	46×10^4	17×10^6	-	0.44	2.7

Średnica grafu
 vs
 ilość wyjętych
 węzłów:
 - przypadkowo
 (failure)
 -w kolejności
 stopnia
 (attack)

R. Albert et al.,
 Nature 406 (2000) 378



Struktura sieci – społeczności

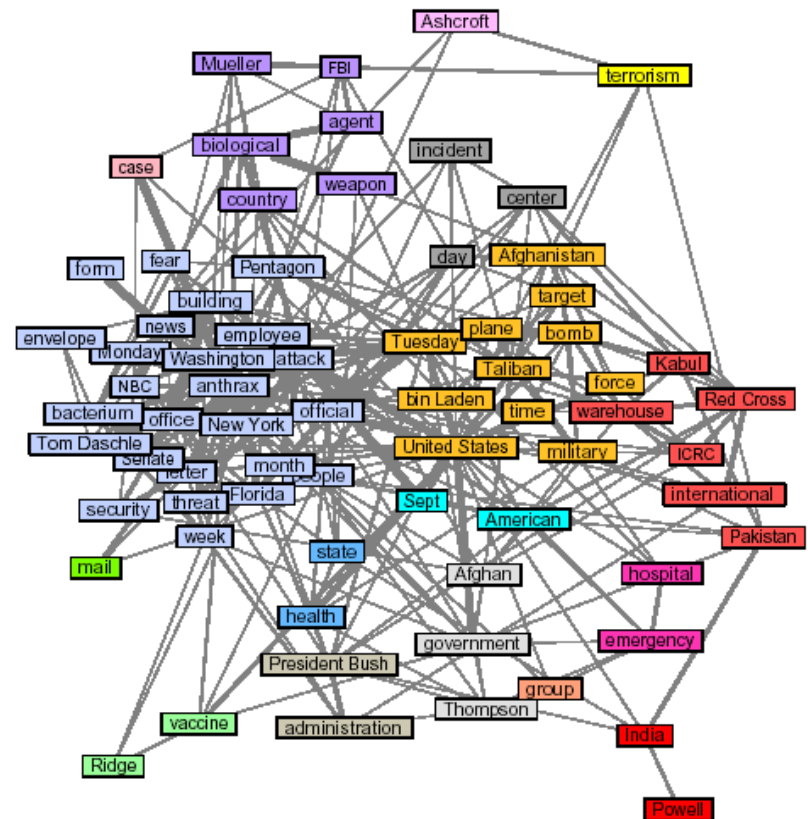
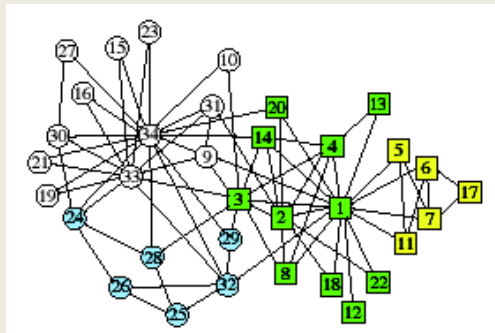
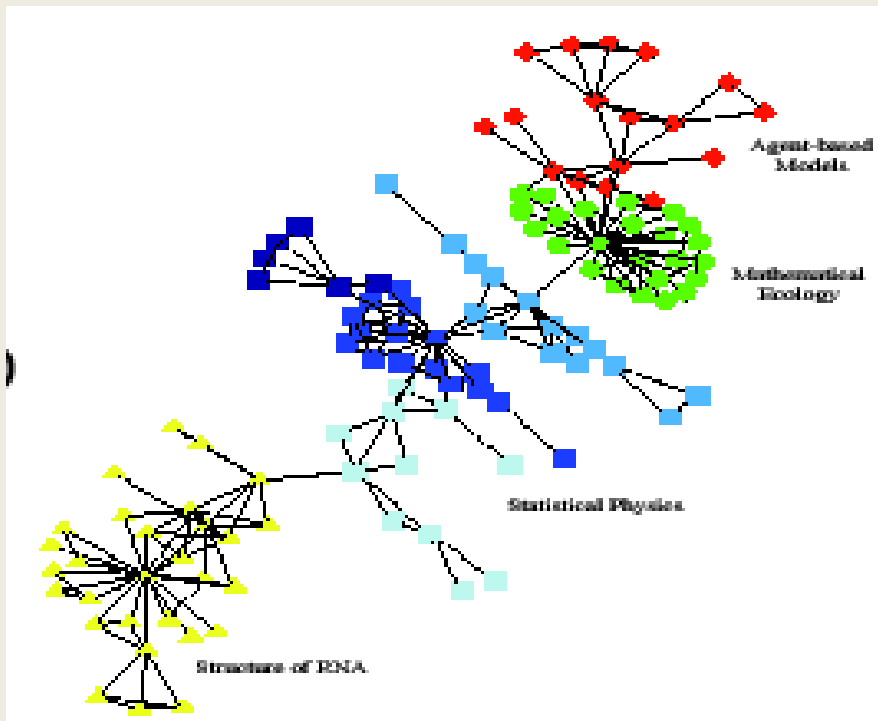


FIG. 3: Network of co-occurrence of words in Reuters newswire stories for October 17, 2001. The widths of the edges indicate their weights and the colors of the vertices indicate the communities found by the algorithm described in the text.

Struktura sieci – społeczności

$$Q = \frac{1}{w} \sum_{ij} \left\{ C_{ij} - \frac{k_i k_j}{w} \right\} \delta(a_i, a_j)$$

gdzie

$$k_i = \sum_{j(i)} C_{ij}$$

$$w = \sum_{ij} C_{ij}$$

Newman, PRE 70 (2004) 56131

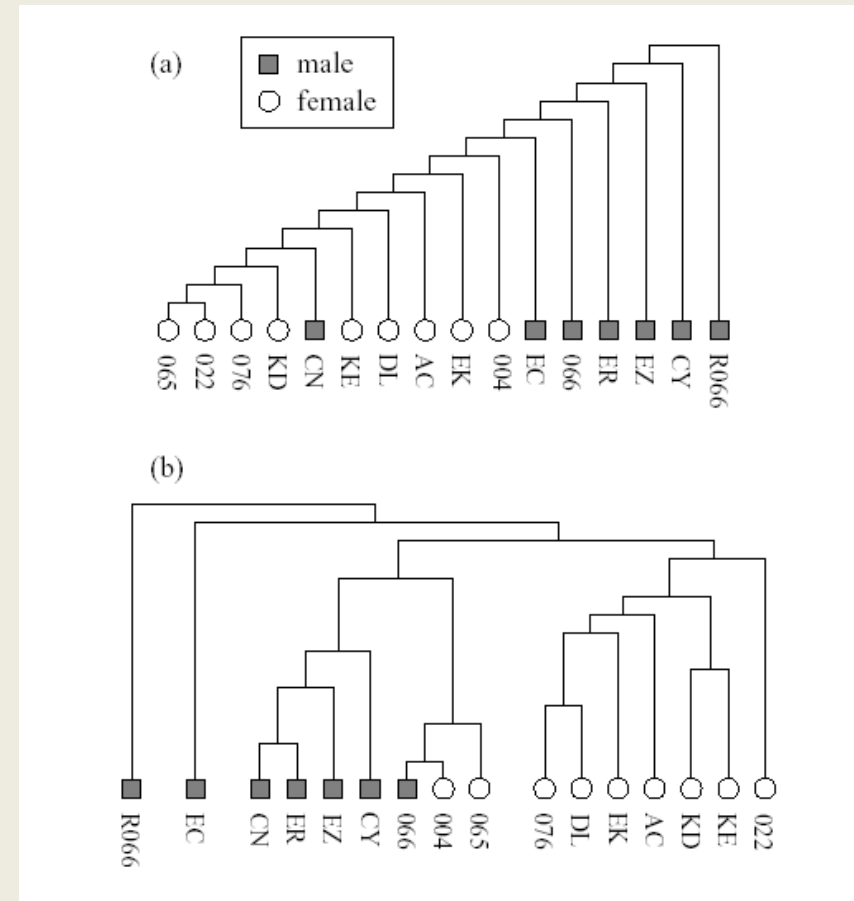


FIG. 2: Community structure in the network of sixteen rhesus monkeys studied by Sade [29]. Squares and circles represent male and female monkeys respectively and the node labels are the same as those used by the original researcher. (a) Dendrogram produced by the algorithm of [27], which ignores the weights on the edges. (b) Dendrogram for the algorithm described here, which takes the weights into account.

Asortatywność stopnia

$$r = \frac{\sum_{jk} jk (e_{jk} - q_j q_k)}{\sigma_q^2}$$

gdzie $q_k = \frac{(k+1)p_{k+1}}{\langle k \rangle}$

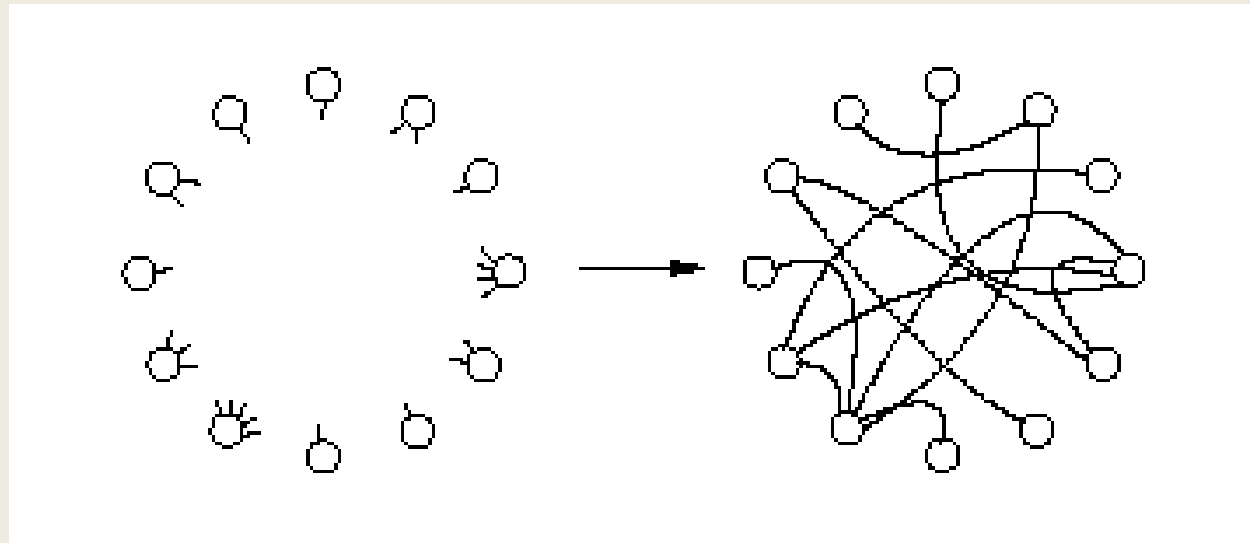
e_{jk} – ułamek wiązań między węzłami o stopniu j, k

	network	type	size n	assortativity r	error σ_r
social	physics coauthorship	undirected	52 909	0.363	0.002
	biology coauthorship	undirected	1 520 251	0.127	0.0004
	mathematics coauthorship	undirected	253 339	0.120	0.002
	film actor collaborations	undirected	449 913	0.208	0.0002
	company directors	undirected	7 673	0.276	0.004
	student relationships	undirected	573	-0.029	0.037
	email address books	directed	16 881	0.092	0.004
technological	power grid	undirected	4 941	-0.003	0.013
	Internet	undirected	10 697	-0.189	0.002
	World-Wide Web	directed	269 504	-0.067	0.0002
	software dependencies	directed	3 162	-0.016	0.020
biological	protein interactions	undirected	2 115	-0.156	0.010
	metabolic network	undirected	765	-0.240	0.007
	neural network	directed	307	-0.226	0.016
	marine food web	directed	134	-0.263	0.037
	freshwater food web	directed	92	-0.326	0.031

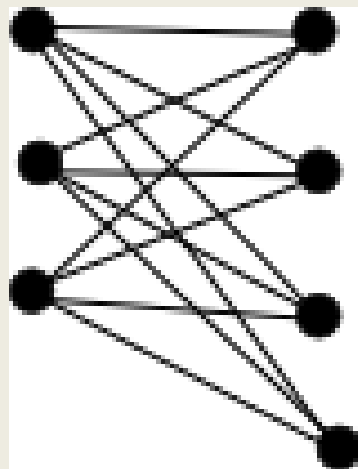
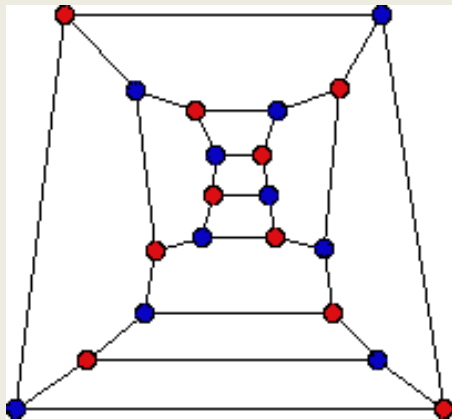


Model konfiguracyjny

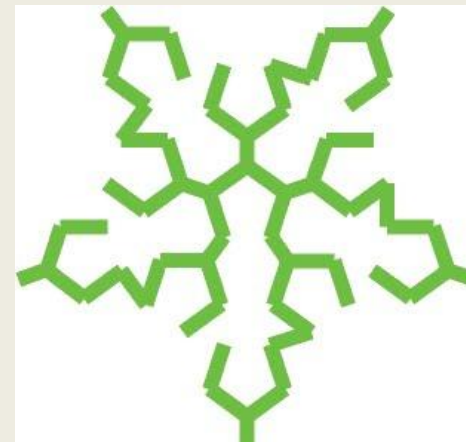
=> dowolny rozkład $P(k)$



Grafy dwudzielne

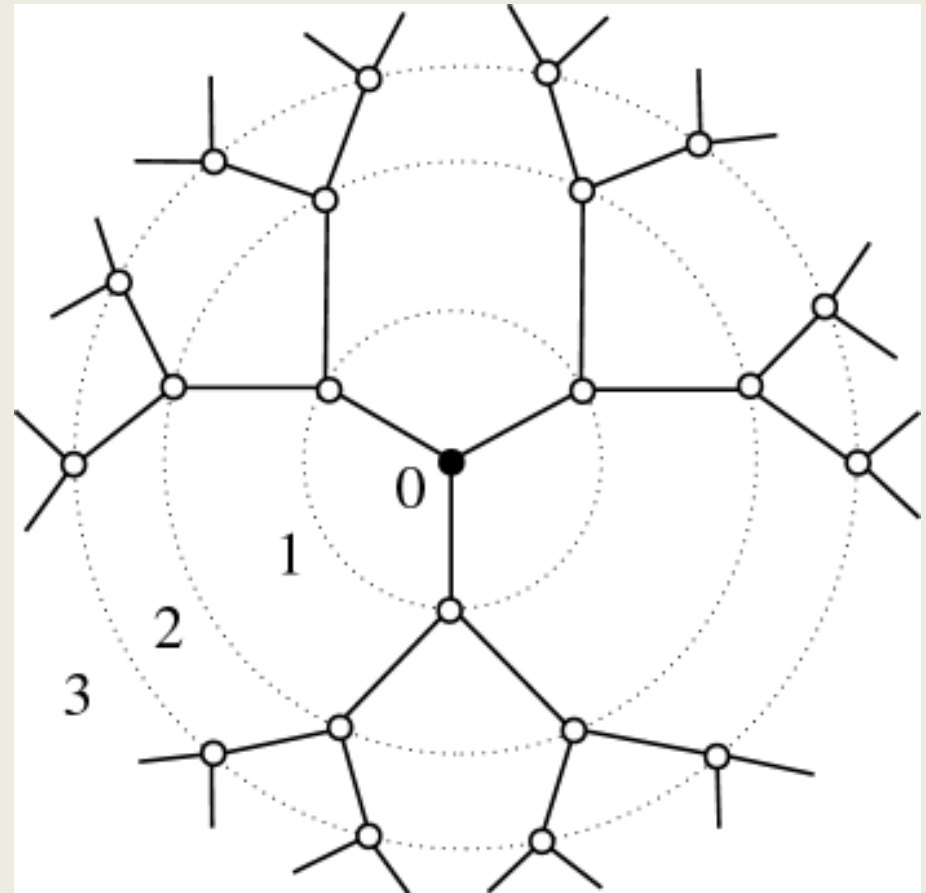
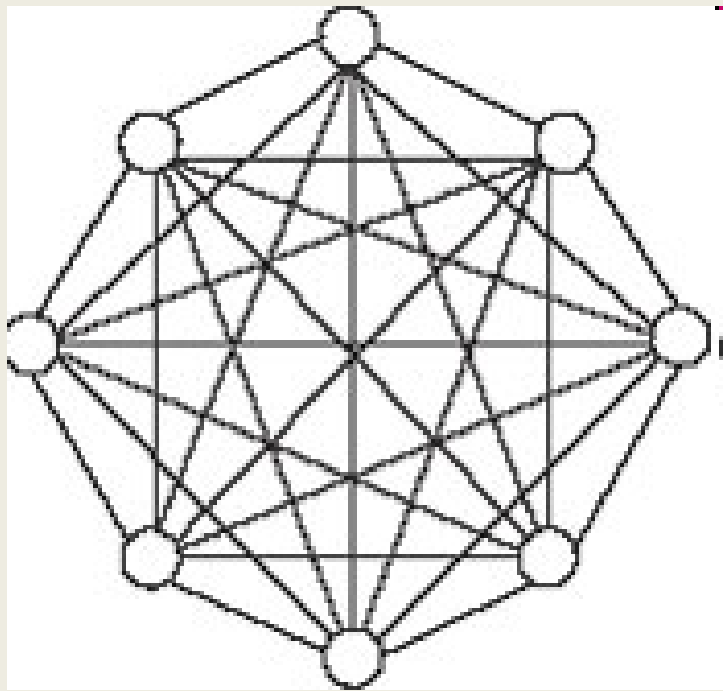


Drzewa



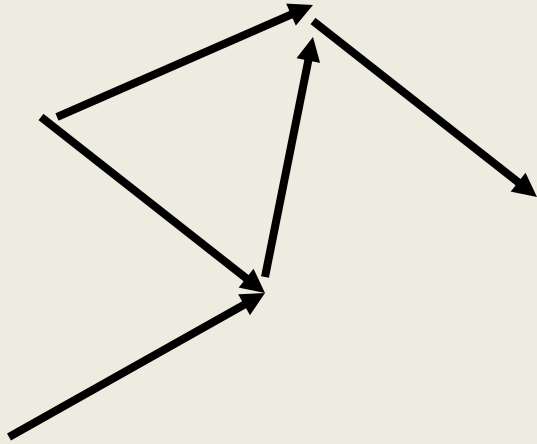
Sieci regularne – każdy węzeł ma tę samą liczbę sąsiadów

Graf w pełni połączony

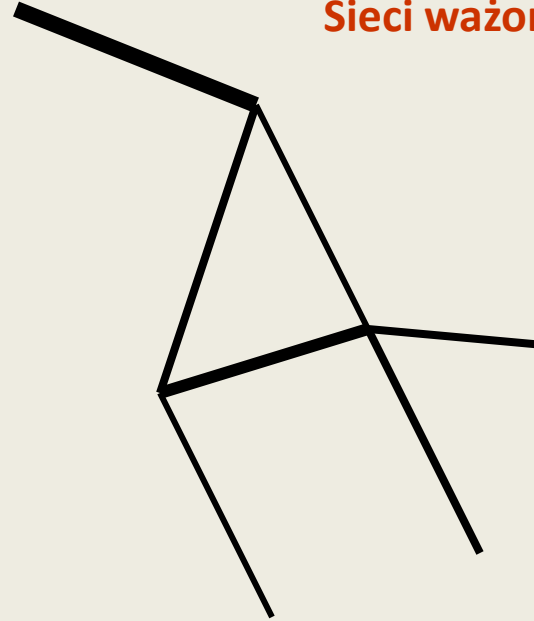


Drzewo Cayleya

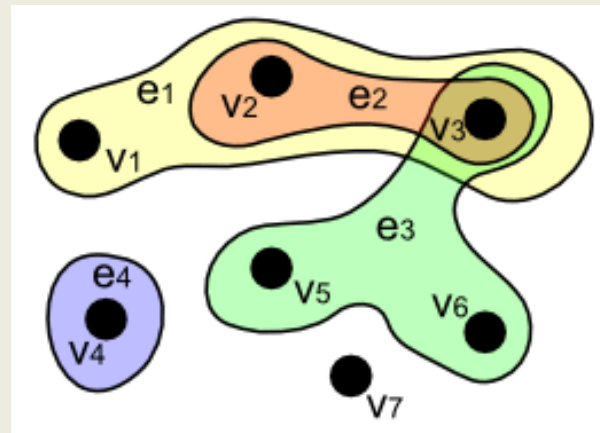
Sieci skierowane



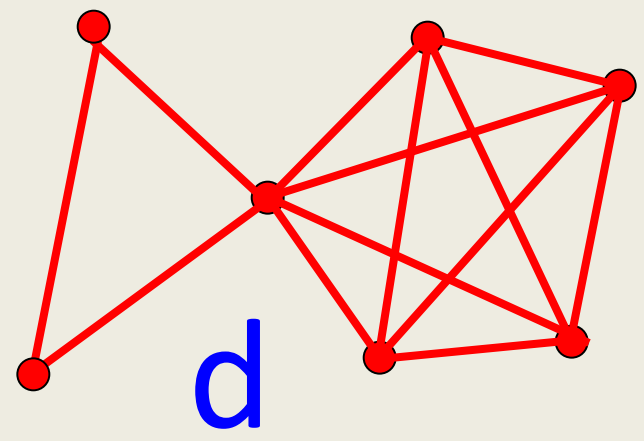
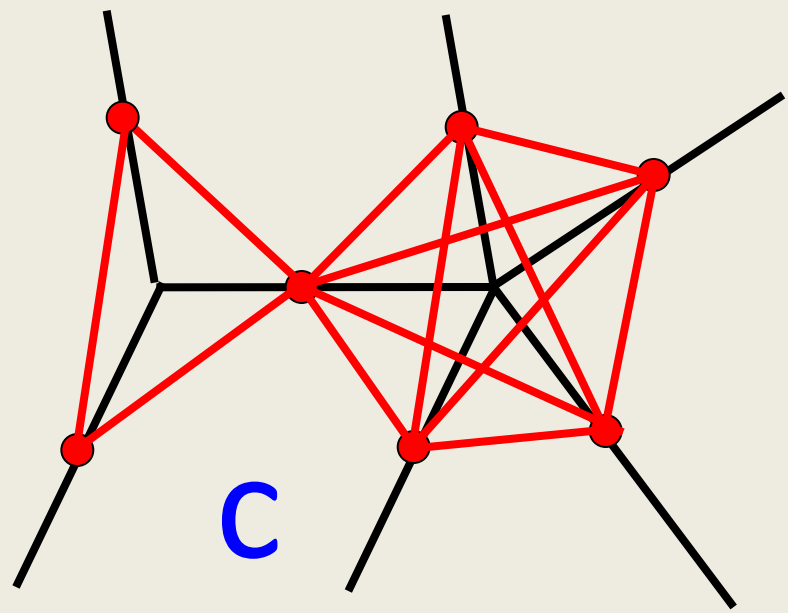
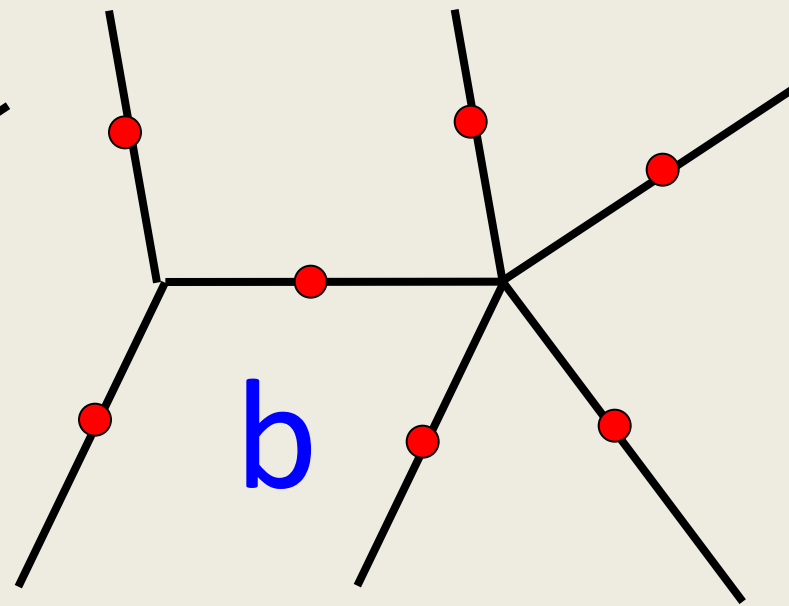
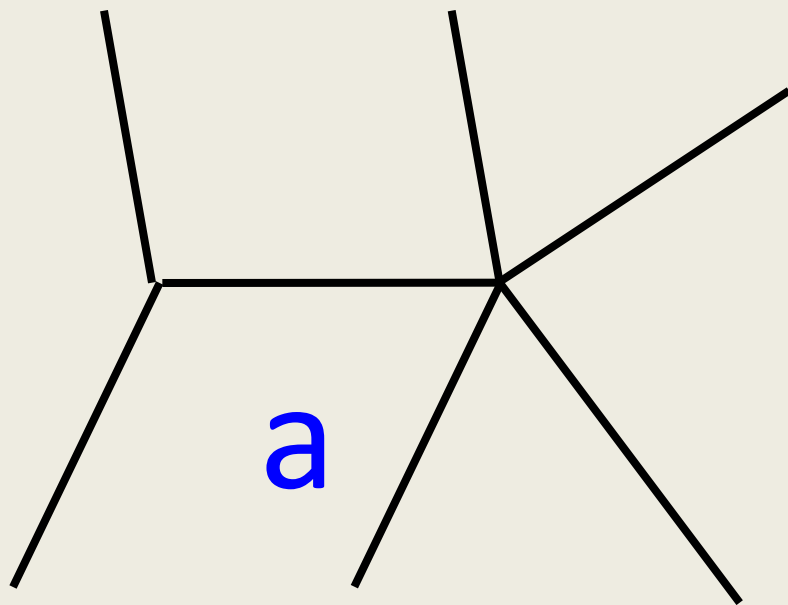
Sieci wazone



Hipergrafy



Graf krawędziowy – konstrukcja



Wybrane problemy badane w zespole (obecnie ZUZ), dotyczące sieci

1. Modelowanie równowagi Heidera w sieciach (PG, PGronek, KK, IJMPC '05)
 2. Efekty pamięci w sieciach rosnących (KM ea, Phys A '05)
 3. Statystyki lawin w sieciach rosnących (KM ea, PRL '05; Phys A '07)
 4. Faza szkła spinowego w sieciach Archimedesesa (MJK, KM ea, PRB '05)
 5. Nowa metoda badania struktury klastrowej sieci (MJK, PRE '08, CPC '10)
 6. Skutki szumu bezskalowego w sieci kwadratowej Isinga (DS, KK, JStatM '08)
 7. Modelowanie współpracy w sieci (PG, KK, Phys A '09)
 8. Widma lawin w sieciach bezskalowych (JTomkowicz, KK, PRE '10)
 9. Struktura grafów krawędziowych (MJK, AM-K, KK, CPC '10, Phys A '11)
 10. Model głosowania na sieciach Kagome i in. (KM, Phys A '11)
 11. Faza szkła spinowego w sieciach jednorodnych (AM-K, doktorat '11)
 12. Model komunikacji i mediów (KM, PGronek, KK, JASSS'11)
 13. Nowa metoda kompresji sieci stanów (MJK, Phys A '11, PRE '12?)
- ...

Why social networks are different from other types of networks

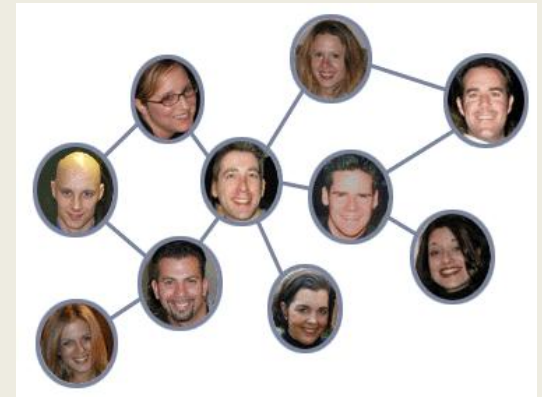


In this paper we have argued that social and non-social networks differ in two important ways. First, they show distinctly different patterns of correlation between the degrees of adjacent vertices, with degrees being positively correlated (**assortative mixing**) in **most social networks** and negatively correlated (disassortative mixing) in most non-social networks. Second, **social networks show high levels of clustering** or network transitivity, whereas clustering in many non-social networks is no higher than one would expect on the basis of pure chance, given the observed degree distribution.

We have shown that both of these differences can be explained by the same hypothesis, that social networks are divided into communities, and non-social networks are not.

Role of line graphs here

In the model proposed here, a social network is the line graph of an initial network of families, communities, interest groups, school classes and small companies. These groups play the role of nodes, and individuals are represented by links between these nodes.



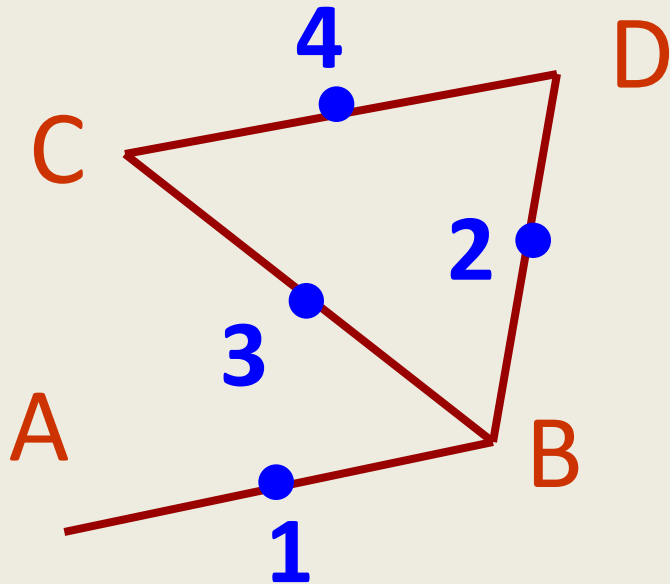
[M J Krawczyk et al, Phys A 2011]



Algorithm

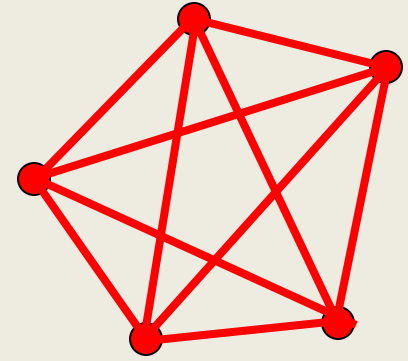
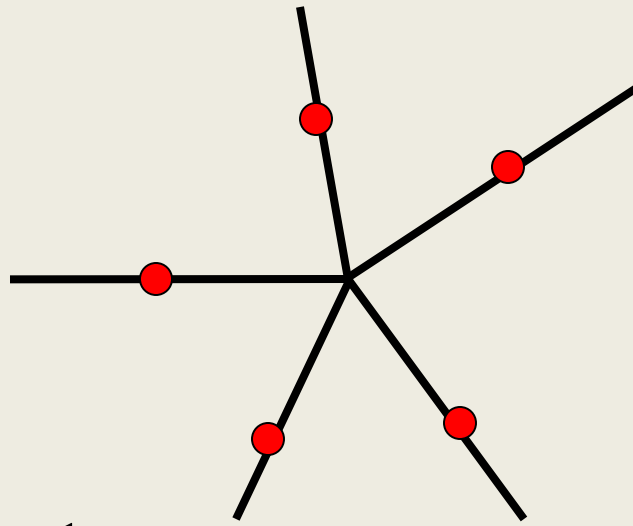
1. assign numbers to the elements of the connectivity matrix above the diagonal. Make the matrix symmetric.

0	1	0	0
1	0	2	3
0	2	0	4
0	3	4	0



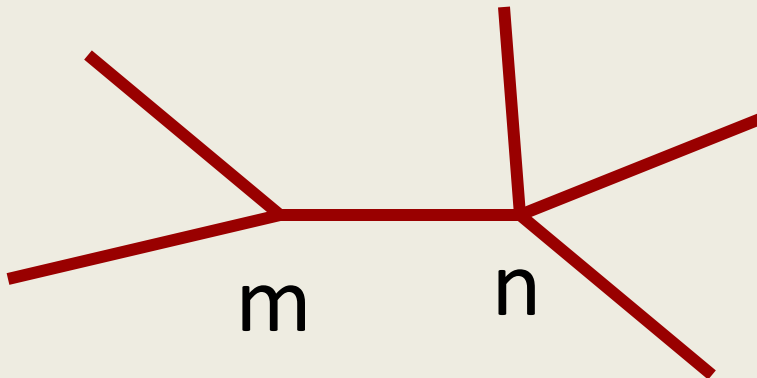
2. If i, j are in the same row or column, then the element $C(i, j)$ of the transformed matrix is 1

0	1	1	0
1	0	1	1
1	1	0	1
0	1	1	0



$$P_t(k) = \frac{1}{\langle k \rangle^2} \sum_{m,n} mP(m)nP(n)\delta_{k,m+n-2}$$

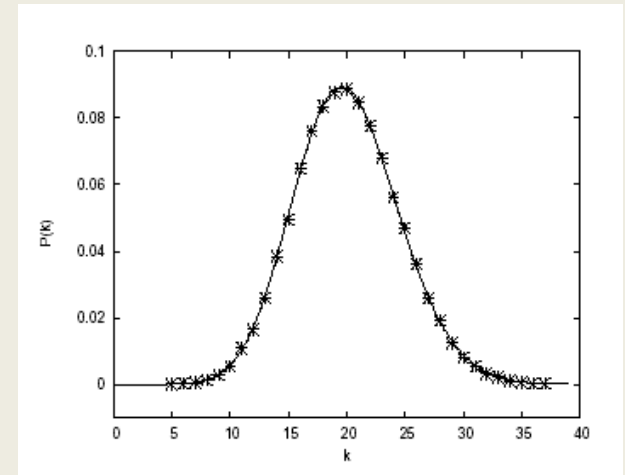
$$= \frac{1}{\langle k \rangle^2} \sum_m mP(m)(k-m+2)P(k-m+2)$$



Degree distribution $P(k)$

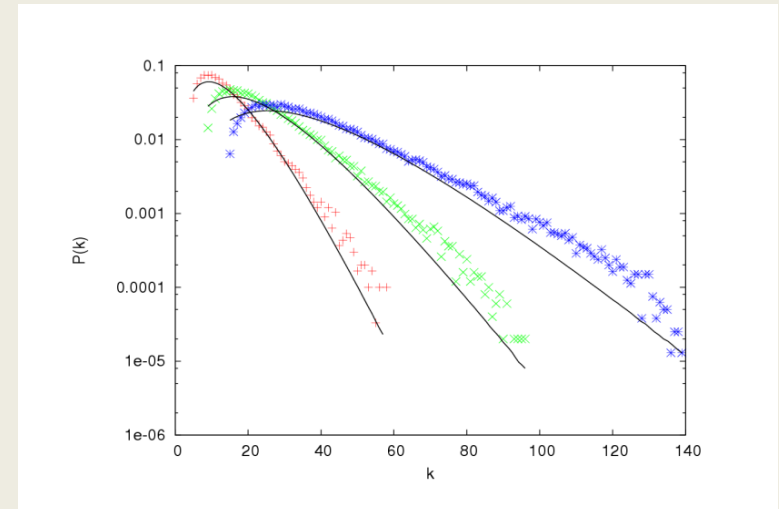
Erdős-Rényi networks

$$P_t(k) = \lambda^k e^{-2\lambda} \sum_{n=1}^{k+1} \frac{1}{(n-1)!(k-n+1)!} =$$
$$= e^{-2\lambda} \frac{(2\lambda)^k}{k!}$$



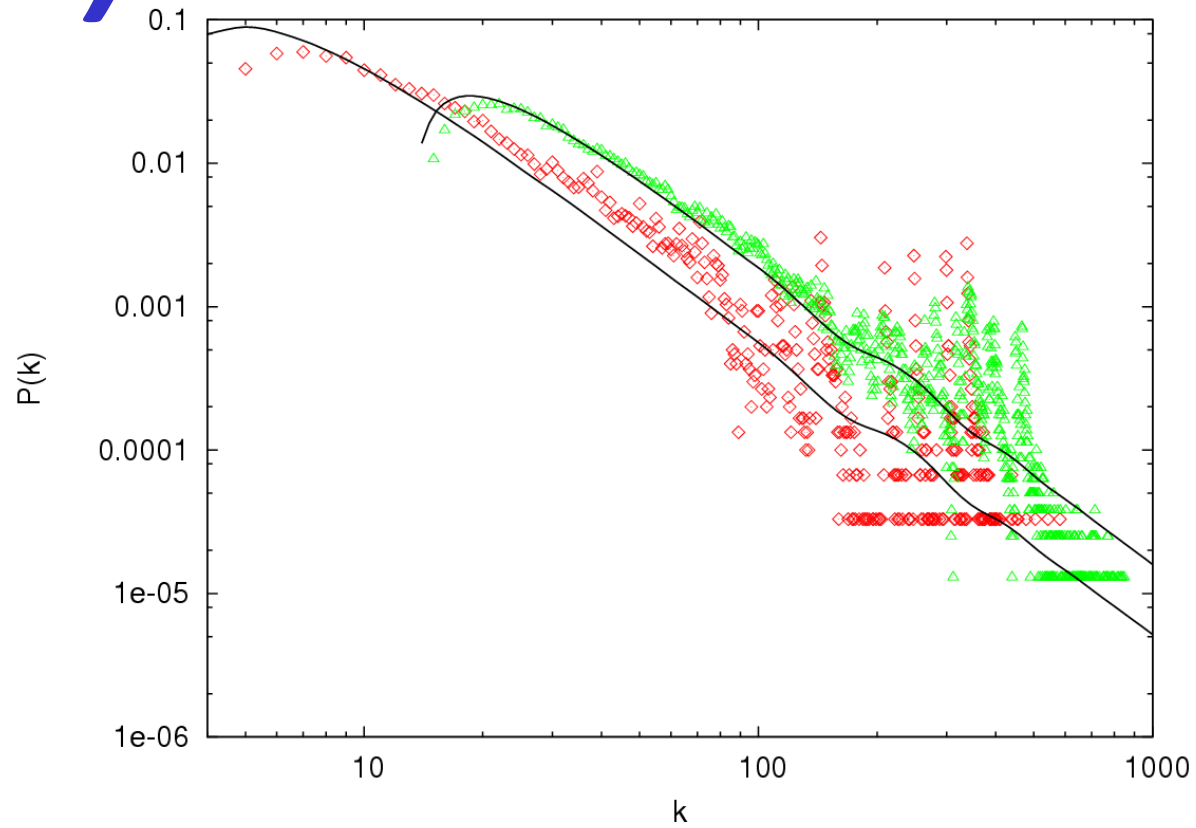
exponential networks

$$P_t(k) = \frac{(1-c)^4}{6} (k+1)(k+2)(k+3)c^k$$



The degree distribution of a line graph on a scale-free network

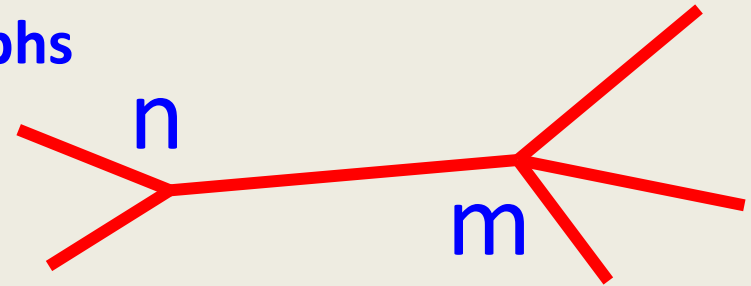
$P_t(k)$



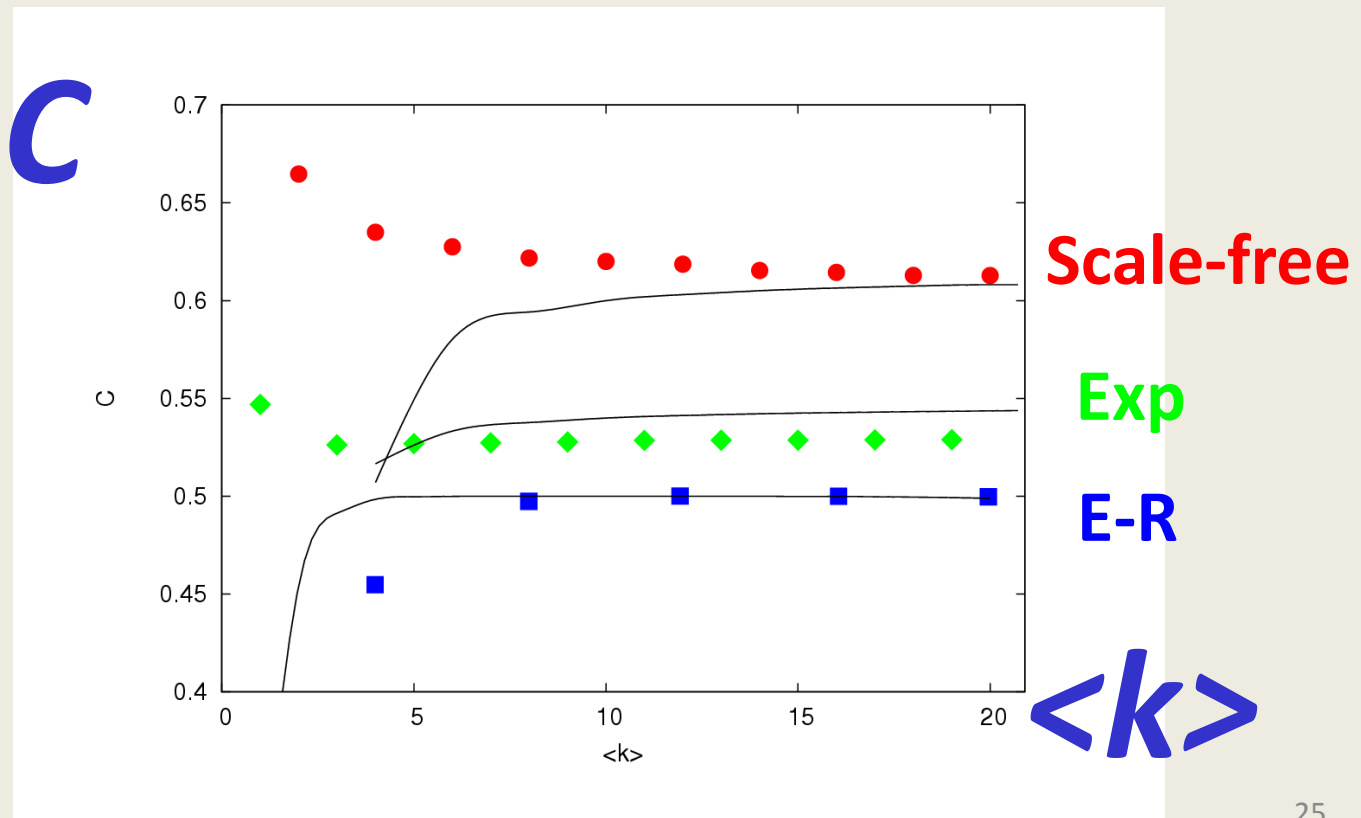
k

$\langle k \rangle = 6, 16$

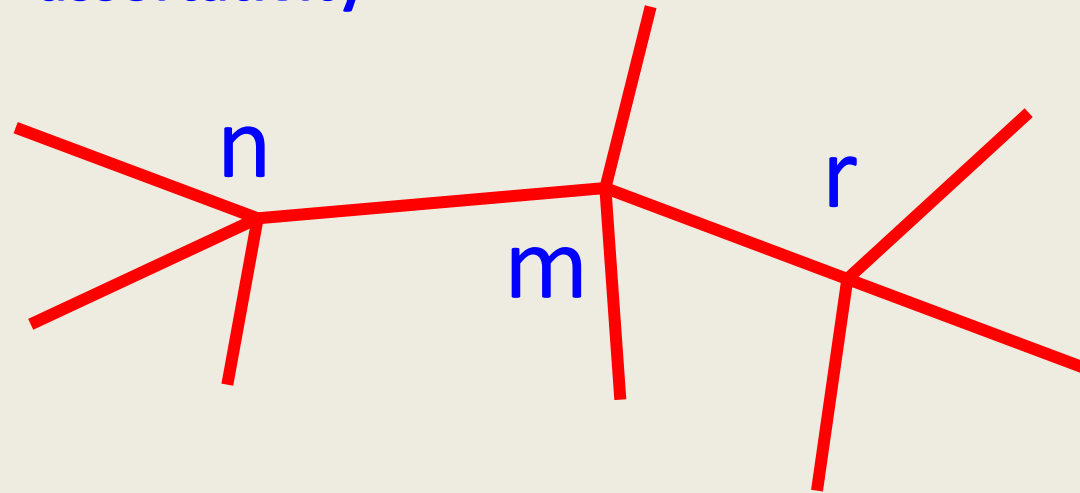
Clustering coefficient in line graphs



$$C = \frac{1}{\lambda^2} \sum_{n=1}^{\infty} nP(n) \sum_{m=1}^{\infty} mP(m) \frac{(n-1)(n-2) + (m-1)(m-2)}{(n+m-2)(n+m-3)}$$



Line graphs - assortativity



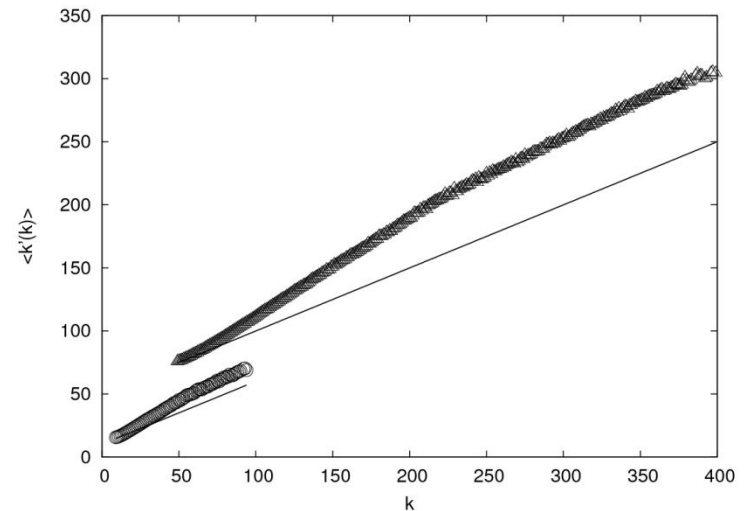
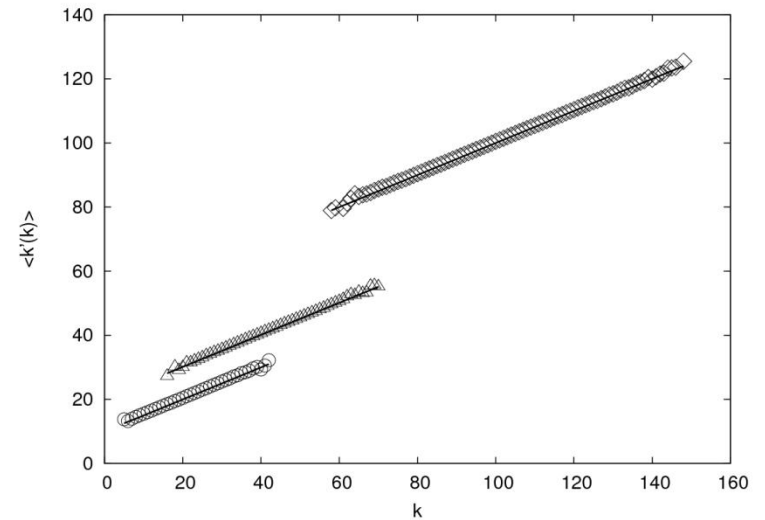
$$\langle k'(k) \rangle = \frac{\sum_n nP(n) \sum_m mP(m) \sum_r rP(r) (n+m-2) \delta_{k, m+r-2}}{\sum_n nP(n) \sum_m mP(m) \sum_r rP(r) \delta_{k, m+r-2}}$$

Assortativity $\langle k'(k) \rangle$

Erdős-Rényi networks

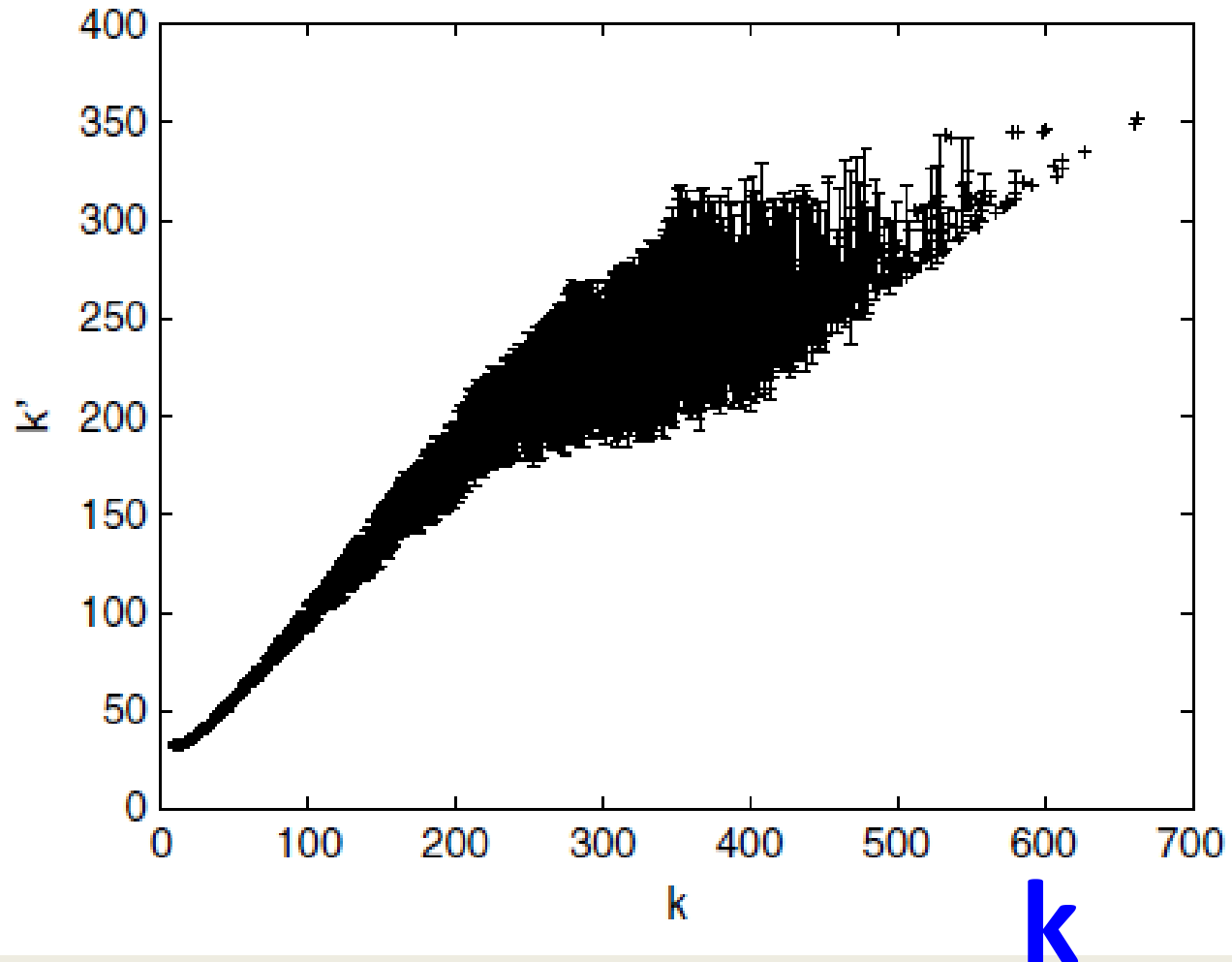
$$\langle k'(k) \rangle = \lambda + \frac{2^{k-1} k}{2^k - 1} \approx \lambda + \frac{k}{2}$$

exponential networks



Assortativity of a line graph on a scale-free network

$$\langle k'(k) \rangle$$



An average over 100 samples of line graphs, constructed for 100 scale-free networks²⁸

LiveJournal

LiveJournal is a remarkably popular platform for personal blog management, populated with over 8 million blogs and over 1 million of communities. LiveJournal was among the first of such platforms available online and it still remains one of the most active and popular. Its users manage personal blogs where they share their daily experiences, political views or discuss news events. Users can also comment on posts of other users.



We defined the network nodes to correspond to personal blogs. Directional links connecting these nodes represent the record that a particular user (owning one blog) monitors another blog (owned by another user).

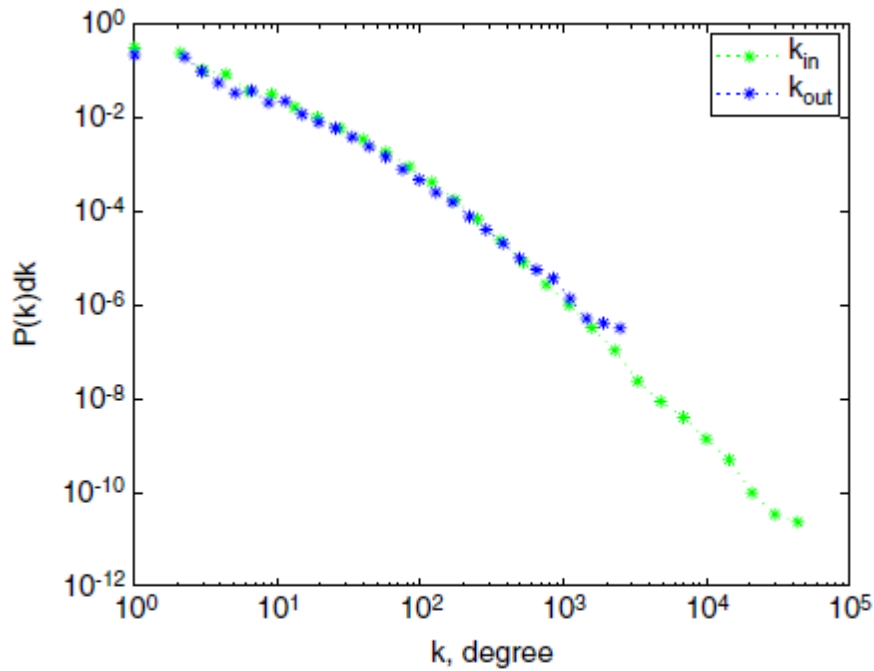


Fig. 2. The degree distribution of the social network of LiveJournal.

**The degree distribution
for LiveJournal
and for the reconstructed network**

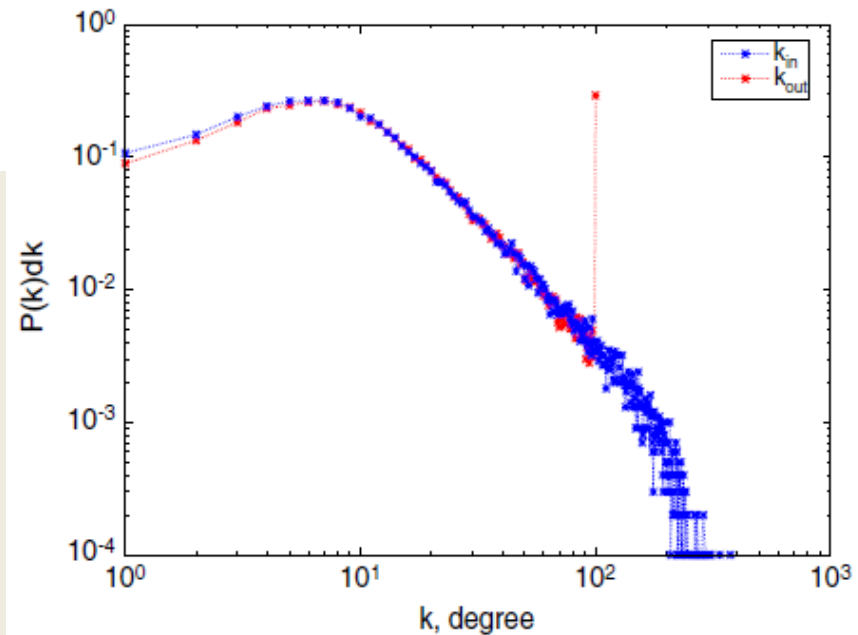
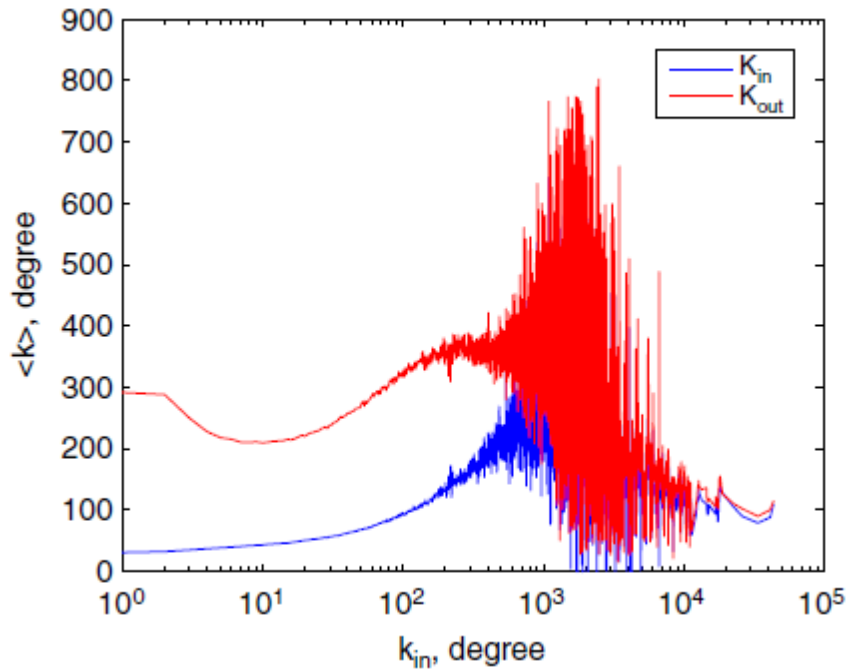
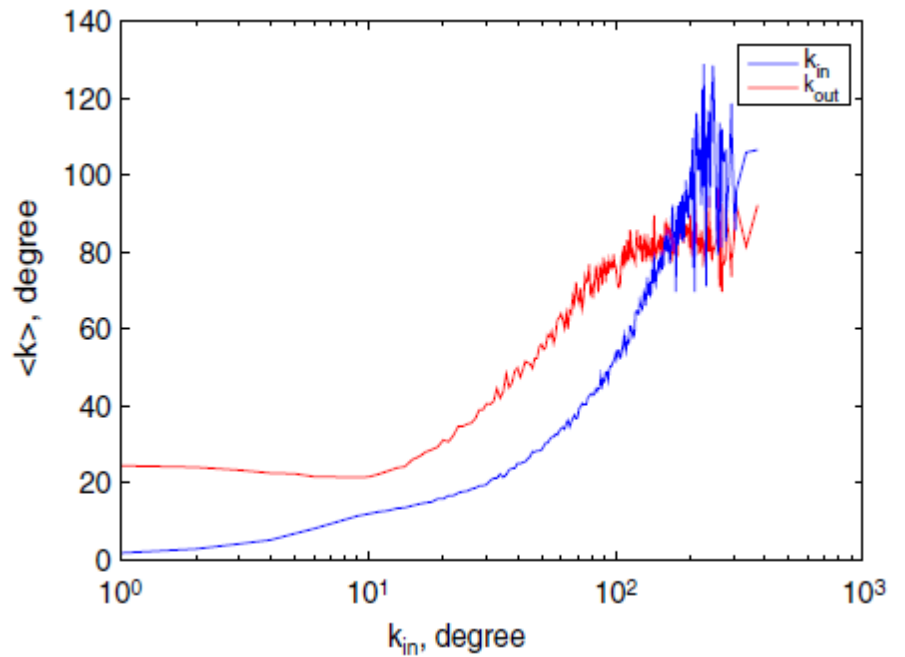
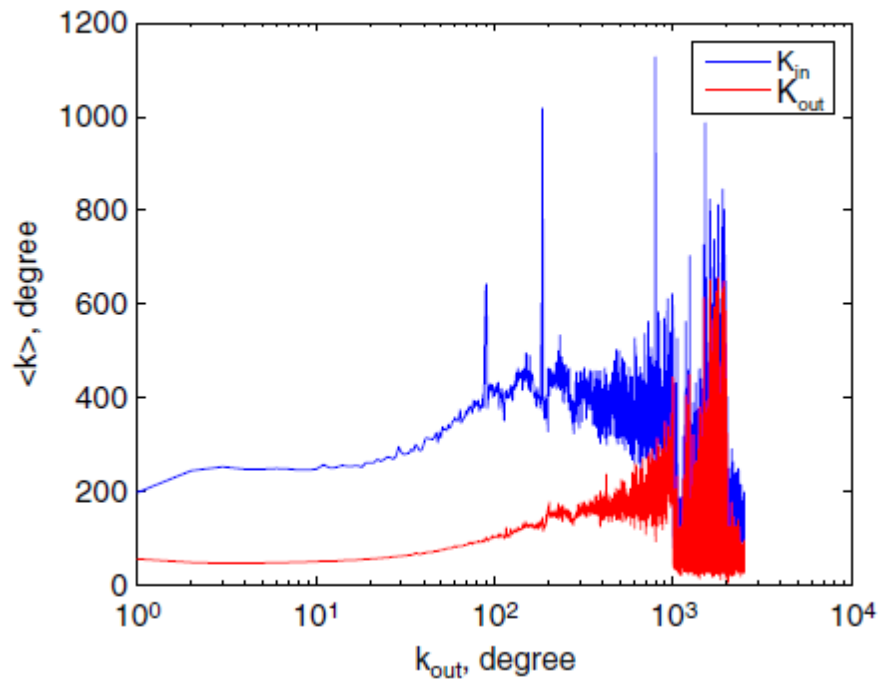


Fig. 6. The degree distribution for the reconstructed network.

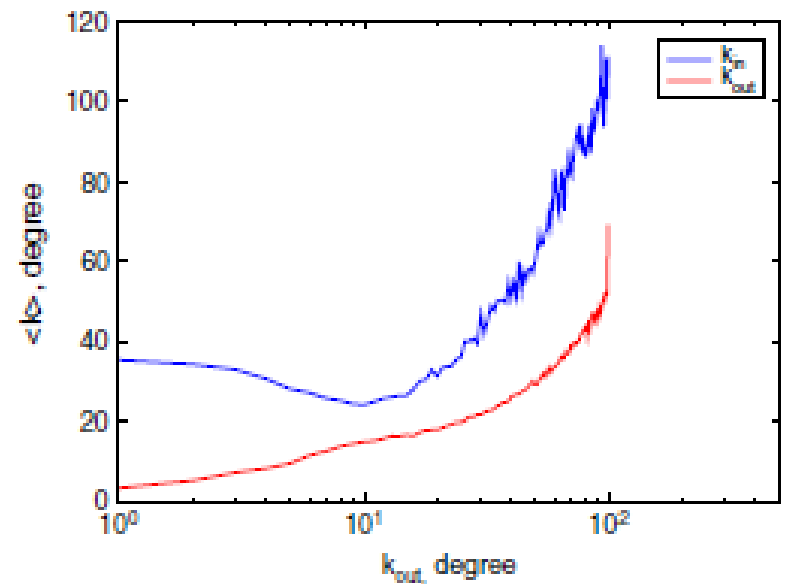


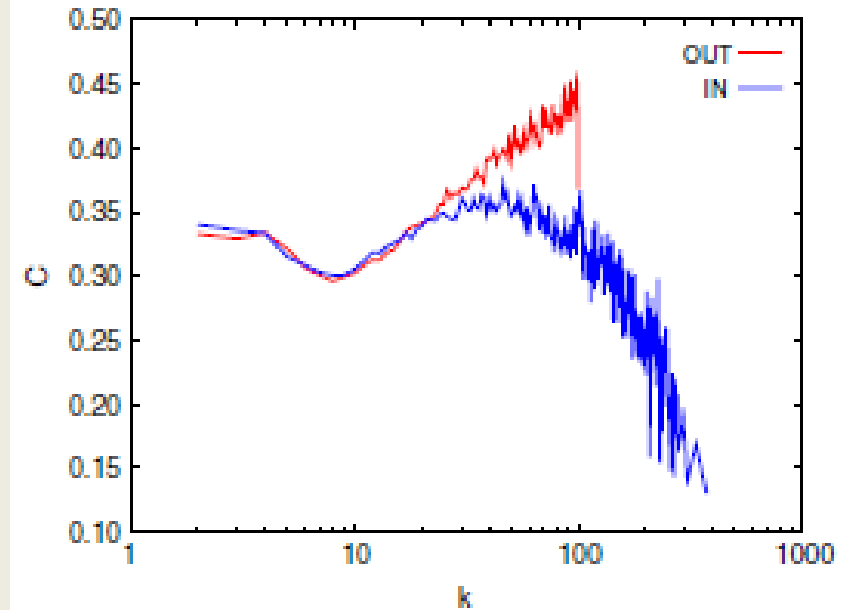
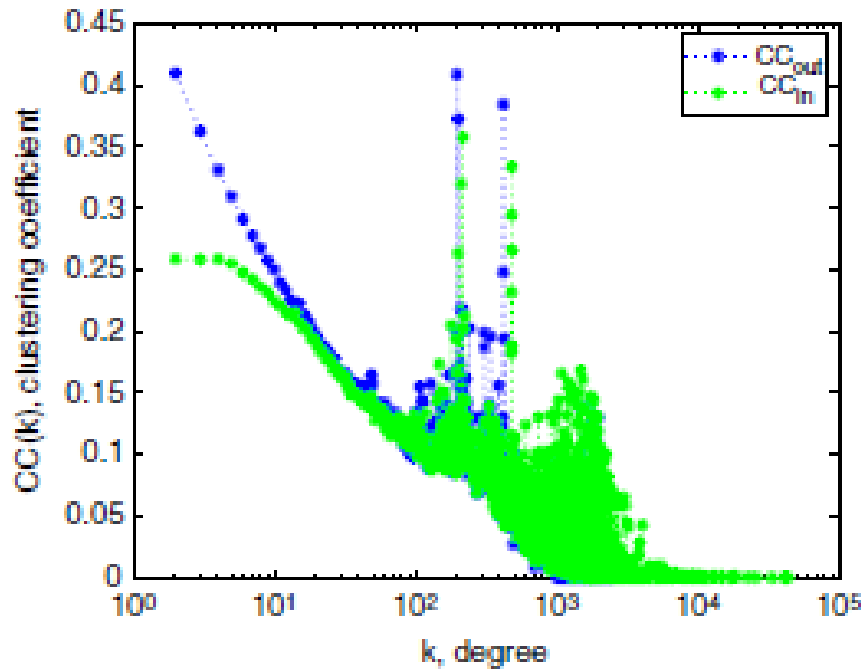
**Mean in-degree and out-degree of
in-neighbours
for LiveJournal
and for the reconstructed network**





**Mean in-degree and out-degree of
out-neighbours
for LiveJournal
and for the reconstructed network**





The degree dependence of the clustering coefficient C for LiveJournal and for the reconstructed network

[M.E.J.Newman, SIAM Review 2003]

