

Omofilia e Assortative Mixing nelle reti

Letterio Galletta

16/05/2012

- Un individuo tende a formare legami sociali con chi è più **simile** a lui (**omofilia**)
- Il comportamento di un individuo tende a essere **simile** a quello dei suoi vicini

Omofilia

- **Limita** il “mondo sociale” di un individuo influenzandone le attitudini e le esperienze
- La **distanza** in termini di “caratteristiche sociali” implica distanza nel grafo delle relazioni sociali

I primi studi sistematici (**sociologia**)

Consideravano piccoli gruppi di persone

- Bambini nelle scuole
- Studenti universitari
- Vicini di casa

Risultati

Tendenza a formare legami con persone demograficamente (età, sesso, razza) e psicologicamente (intelligenza, attitudini, aspirazioni, educazione) simili

Caratteristiche socio-demografiche

- Razza/Origini etniche
- Sesso/Genere
- Età
- Religione
- Educazione
- Classe sociale

Sono fonte di omofilia e **strutturano** la società

Razza/Origini etniche

- Sono una piccola parte delle persone considerano “**confidenti**” persone di un'altra razza/gruppo etnico
- Questo fenomeno è stato anche riscontrato in reti sociali piccole (**scuole, aziende**)

Sesso/Genere

Bambini

- Tendono a formare gruppi omogenei
- Sono portati a risolvere l'intransitività “eliminando” un legame cross-sex

Adulti

- Tendono a formare gruppi eterogenei
- Sono portati a risolvere l'intransitività “aggiungendo” un legame cross-sex

Età

- Tendenza ad avere legami con persone della propria età
- Nei bambini delle elementari è più forte
- Frischer (1977) su uomini di Detroit
 - 38% - 2 anni
 - 72% - 8 anni

La sociologia ha mostrato che gli individui tendono ad aggregarsi in gruppi **omogenei**

Domande:

- È possibile dare una misura dell'omofilia in una rete?
- L'omofilia è un fenomeno che si manifesta solo nelle reti sociali?

Assortative mixing in networks

Assortative mixing \approx omofilia in rete generiche

M. E. J. Newman. Mixing patterns in networks

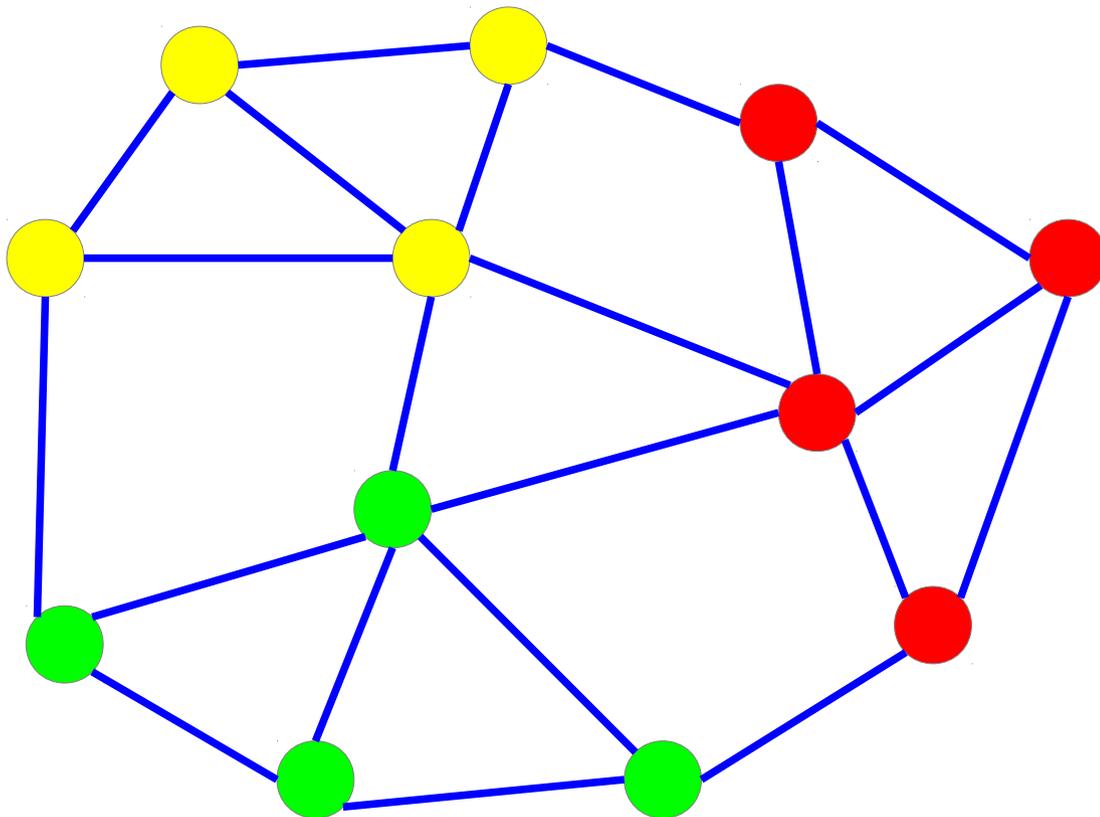
- Metodo per misurare il livello di assortative mixing di una rete
- Misurazioni su diversi tipi di rete
- Modelli e algoritmi di simulazione

Assortative mixing

- I nodi di una rete tendono ad avere archi con nodi che presentano caratteristiche simili
- Due tipi di assortative mixing
 - Discreta => i nodi sono classificati in base alle loro caratteristiche intrinseche (**tipo**)
 - Scalare => i nodi sono classificati in base a delle proprietà scalari (**grado**)

Assortative mixing discreto

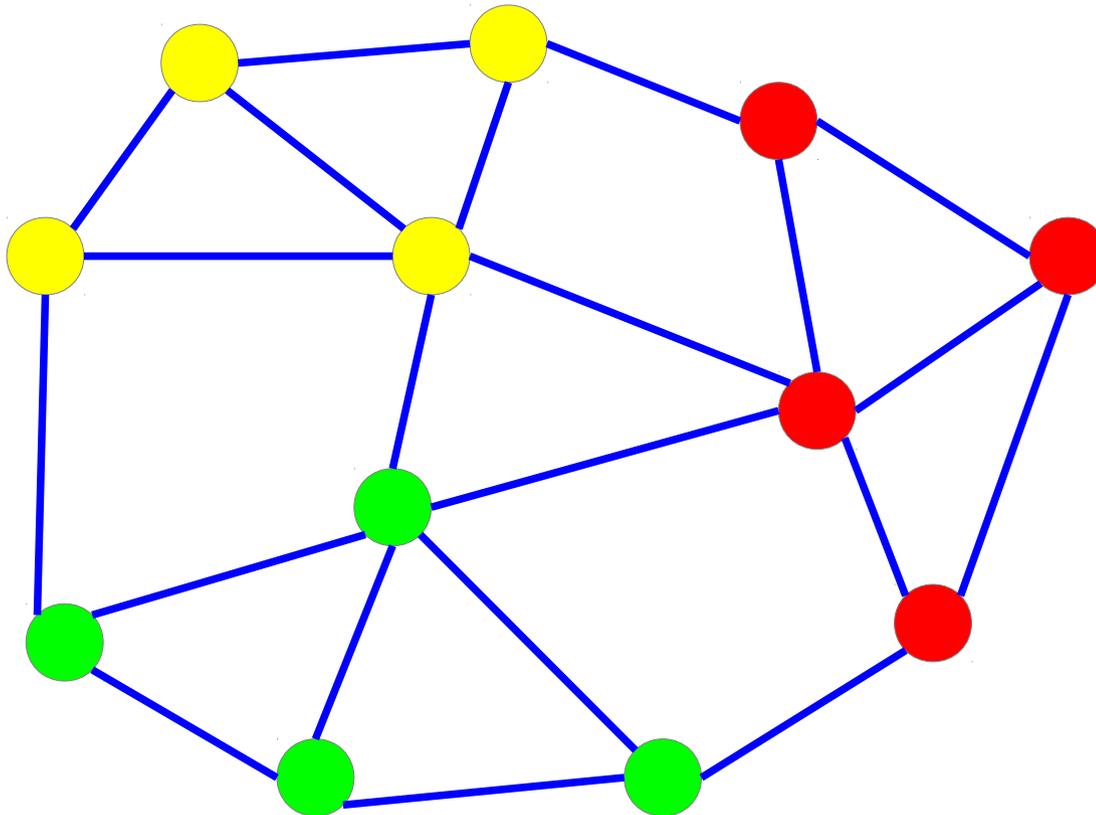
Una rete dove i nodi sono caratterizzati da un tipo



- 12 nodi
- 21 archi
- 3 tipi

Assortative mixing discreto

Calcoliamo e_{ij} la frazione degli archi della rete che connettono nodi di tipo i e di tipo j



e_{ij}	G	V	R
G	5/21	2/21	2/21
V	2/21	5/21	2/21
R	2/21	2/21	5/21

Assortative mixing discreto

- Calcoliamo a_i la frazione degli archi uscenti da nodi di tipo i
- Calcoliamo b_j la frazione degli archi entranti in nodi di tipo j

e_{ij}	G	V	R	a
G	5/21	2/21	2/21	3/7
V	2/21	5/21	2/21	3/7
R	2/21	2/21	5/21	3/7
b	3/7	3/7	3/7	

$$a_i = \sum_j e_{ij}$$

$$b_j = \sum_i e_{ij}$$

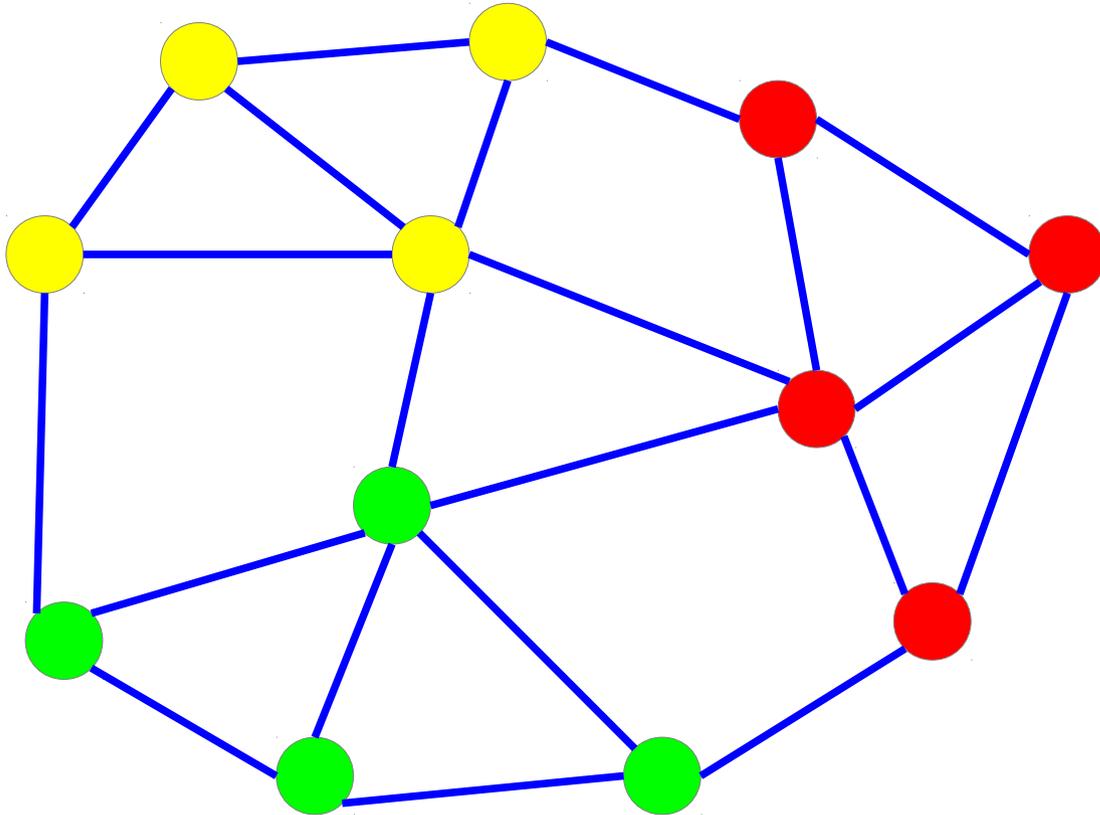
Coefficiente di assortativity

$$R = \frac{\sum_i e_{ii} - \sum_i a_i \cdot b_i}{1 - \sum_i a_i \cdot b_i}$$

Un numero reale compreso tra -1 e 1:

- $R = 0$ nessun assortative mixing
- $R = 1$ assortative mixing perfetto
- $R < 0$ disassortative mixing

Nel nostro esempio



e_{ij}	G	V	R	a
G	5/21	2/21	2/21	3/7
V	2/21	5/21	2/21	3/7
R	2/21	2/21	5/21	3/7
b	3/7	3/7	3/7	

$$R = 0.3636$$

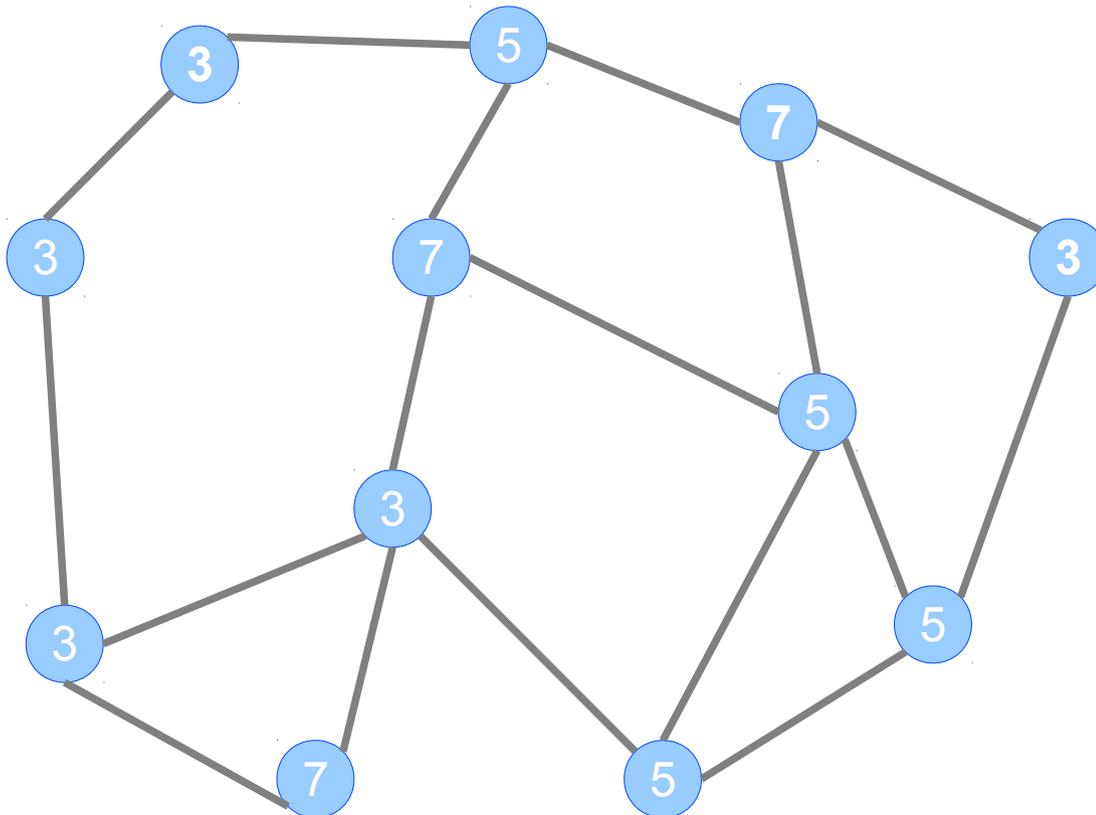
		women				a_i
		black	hispanic	white	other	
men	black	0.258	0.016	0.035	0.013	0.323
	hispanic	0.012	0.157	0.058	0.019	0.247
	white	0.013	0.023	0.306	0.035	0.377
	other	0.005	0.007	0.024	0.016	0.053
b_i		0.289	0.204	0.423	0.084	

TABLE I: The mixing matrix e_{ij} and the values of a_i and b_i for sexual partnerships in the study of Catania *et al.* [23]. After Morris [24].

$$R = 0.621$$

Assortative mixing scalare

Una rete dove i nodi sono caratterizzati da una proprietà scalare



- 12 nodi
- 17 archi
- 3, 5, 7 valori

Assortative mixing scalare

- e_{xy} la frazione degli archi della rete che connettono nodi con valore x a nodi con valore y
- a_x la frazione degli archi uscenti da nodi con valore x
- b_y la frazione degli archi entranti in nodi con valore y

e_{xy}	3	5	7	a
3	3/17	3/17	4/17	10/17
5	3/17	3/17	4/17	10/17
7	4/17	4/17	0	8/17
b	10/17	10/17	8/17	

$$a_x = \sum_y e_{xy}$$

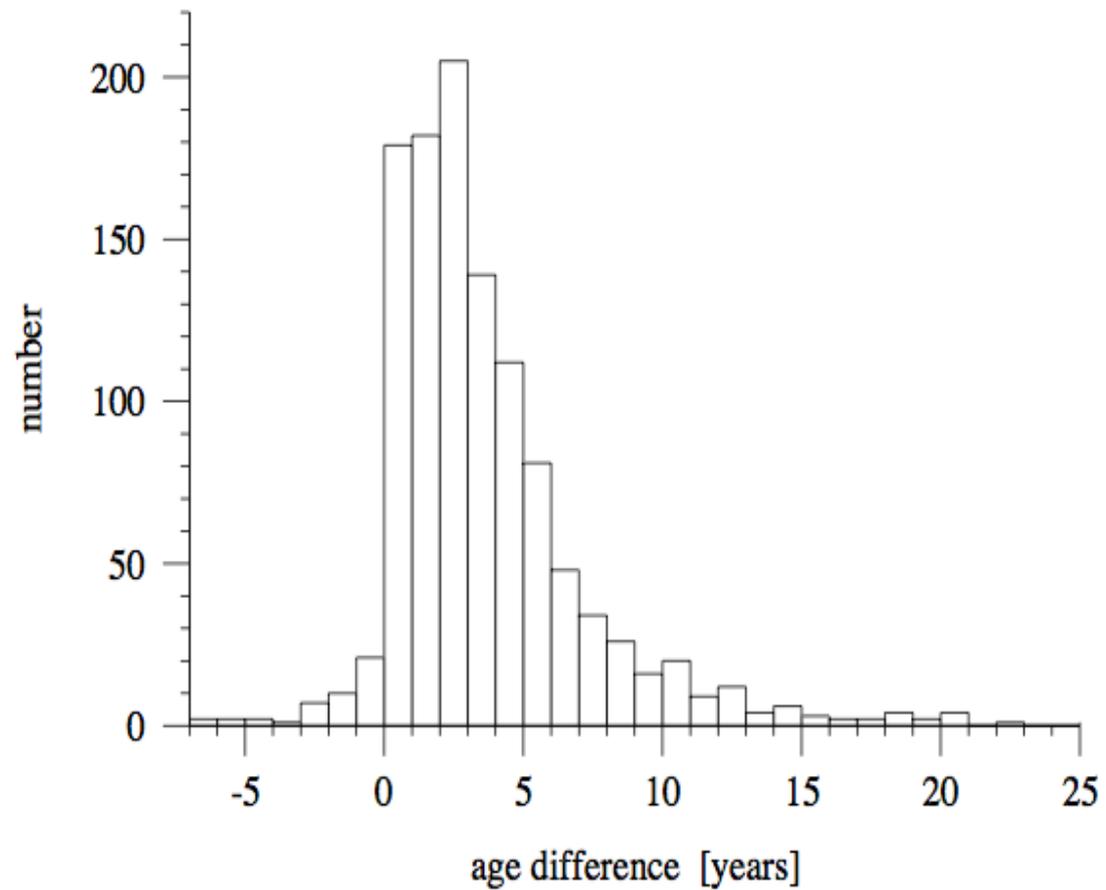
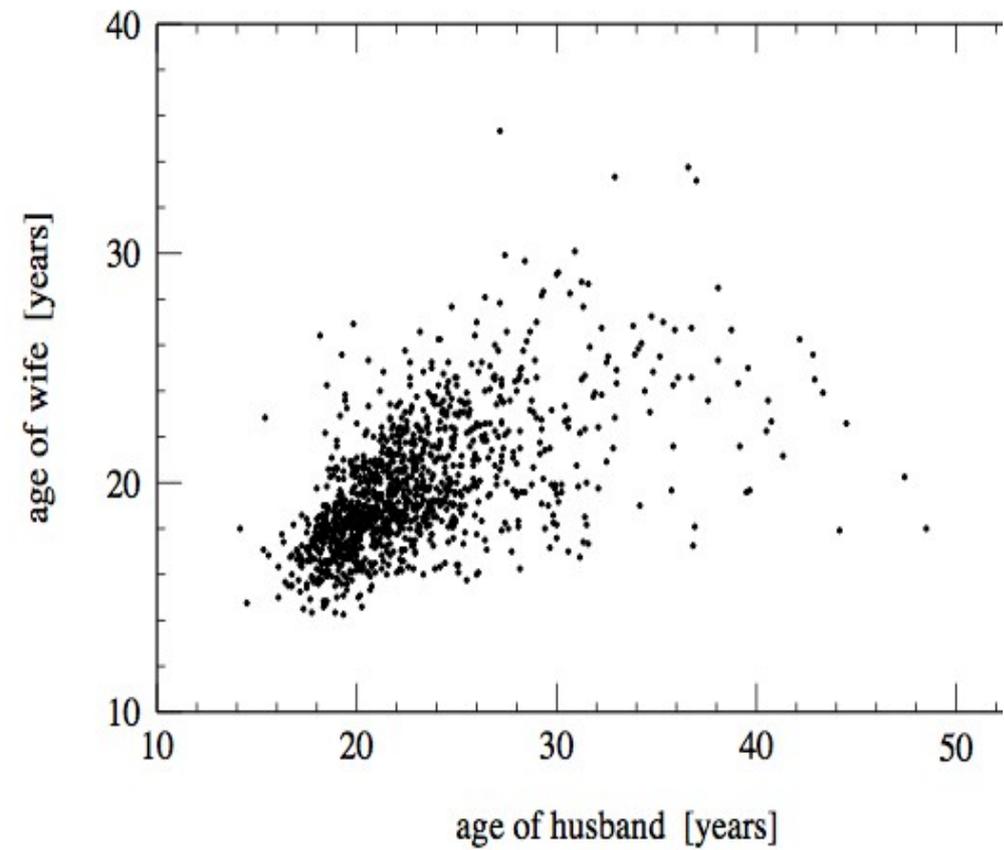
$$b_y = \sum_x e_{xy}$$

Coefficiente di assortativity

$$R = \frac{\sum_{xy} x y (e_{xy} - a_x \cdot b_y)}{\sigma_a \sigma_b}$$

Un numero reale nell'intervallo $[-1, 1]$:

- $R = 0$ nessun assortative mixing
- $R = 1$ assortative mixing perfetto
- $R = -1$ disassortative mixing perfetto



$R = 0.547$

Il grado come proprietà scalare

- Non consideriamo il grado ma il “remaining degree”

$$q_k = \frac{(k+1) p_{k+1}}{\sum_j j p_j}$$

- e_{jk} frazione degli archi che connettono un nodo di remaining degree j a uno con remaining degree k
- Valgono

$$q_k = \sum_j e_{jk} \qquad q_j = \sum_k e_{jk}$$

Coefficiente nel caso del grado

Nel caso delle reti non orientate

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

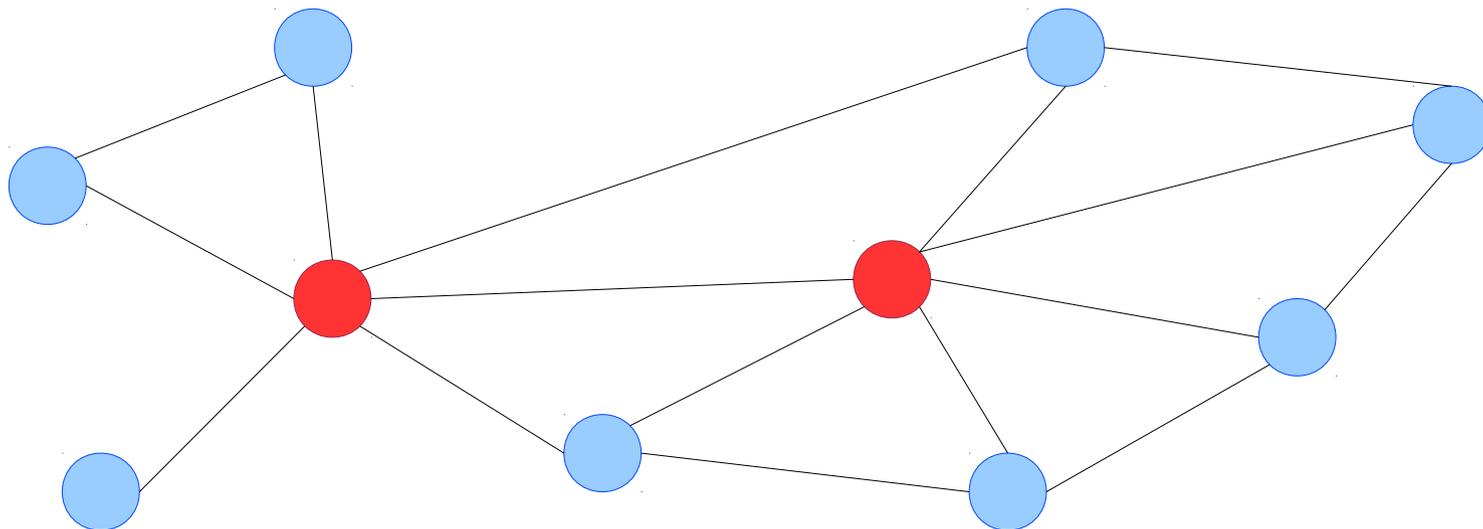
Nel caso delle reti orientate

$$R = \frac{\sum_{jk} jk (e_{jk} - q_j^i q_k^o)}{\sigma_i \sigma_o}$$

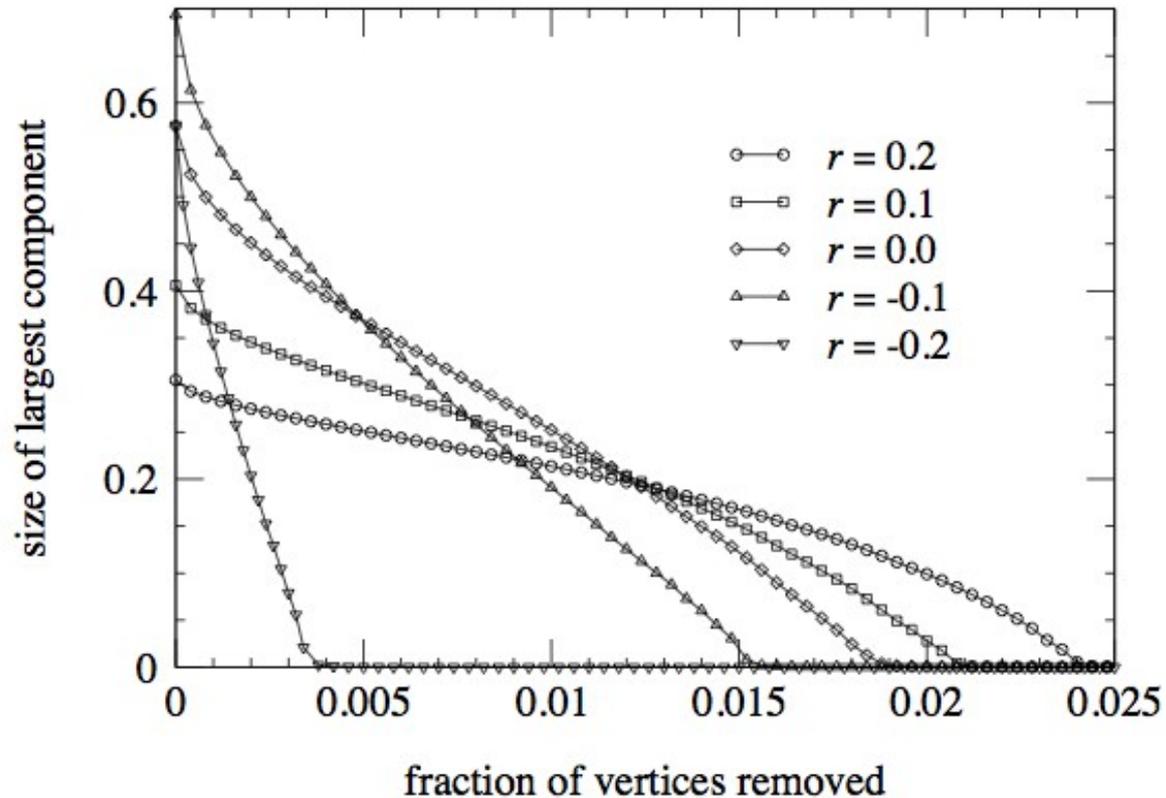
	network	type	size n	assortativity r	error σ_r	ref.
social	physics coauthorship	undirected	52 909	0.363	0.002	a
	biology coauthorship	undirected	1 520 251	0.127	0.0004	a
	mathematics coauthorship	undirected	253 339	0.120	0.002	b
	film actor collaborations	undirected	449 913	0.208	0.0002	c
	company directors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

Robustezza di una rete

- Quanti nodi devono essere eliminati per disconnettere una rete?
- Scelta dei nodi da eliminare
 - random
 - mirata



Assortativity e robustezza



Le reti con un alto coefficiente di assortativity per il grado sono più robuste

Domande:

- È possibile dare una misura dell'omofilia in una rete? (si)
- L'omofilia è un fenomeno che si manifesta solo nelle reti sociali? (si)

- Un individuo tende a formare legami sociali con chi è più **simile** a lui (**homophily**)
- Il comportamento di un individuo tende a essere **simile** a quello dei suoi vicini

Tendenza di comportamenti simili tra vicini

Domande:

- influenza?
- omofilia?

Studi di Christakis e Fowler

Felicità

- Cluster di gente felice/infelice
- La felicità si estende fino a distanza 3
- La presenza di cluster dipende sia dalla diffusione sia da omofilia

Obesità

- Cluster di gente obesa/non obesa
- L'obesità si estende fino a distanza 3
- La presenza di cluster dipende sia dalla diffusione sia da omofilia

Studi di Christakis e Fowler

Vizio del fumo

- Cluster di fumatori /non fumatori
- Il vizio del fumo si estende fino a distanza 3
- La presenza di cluster dipende sia dalla diffusione sia da omofilia

Link:

- [The Collective Dynamics of Smoking in a Large Social Network](#)
- [The Spread of Obesity in Social Networks](#)

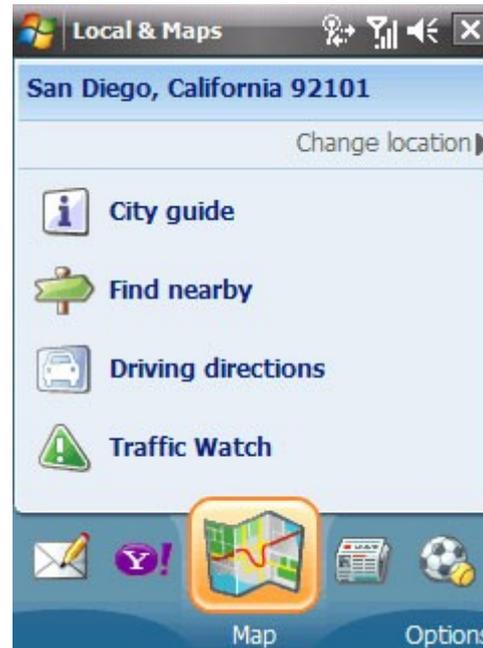
Il problema:

Distinguere gli effetti dell'omofilia e dell'influenza dei vicini nella diffusione dei comportamenti

Una possibile soluzione:

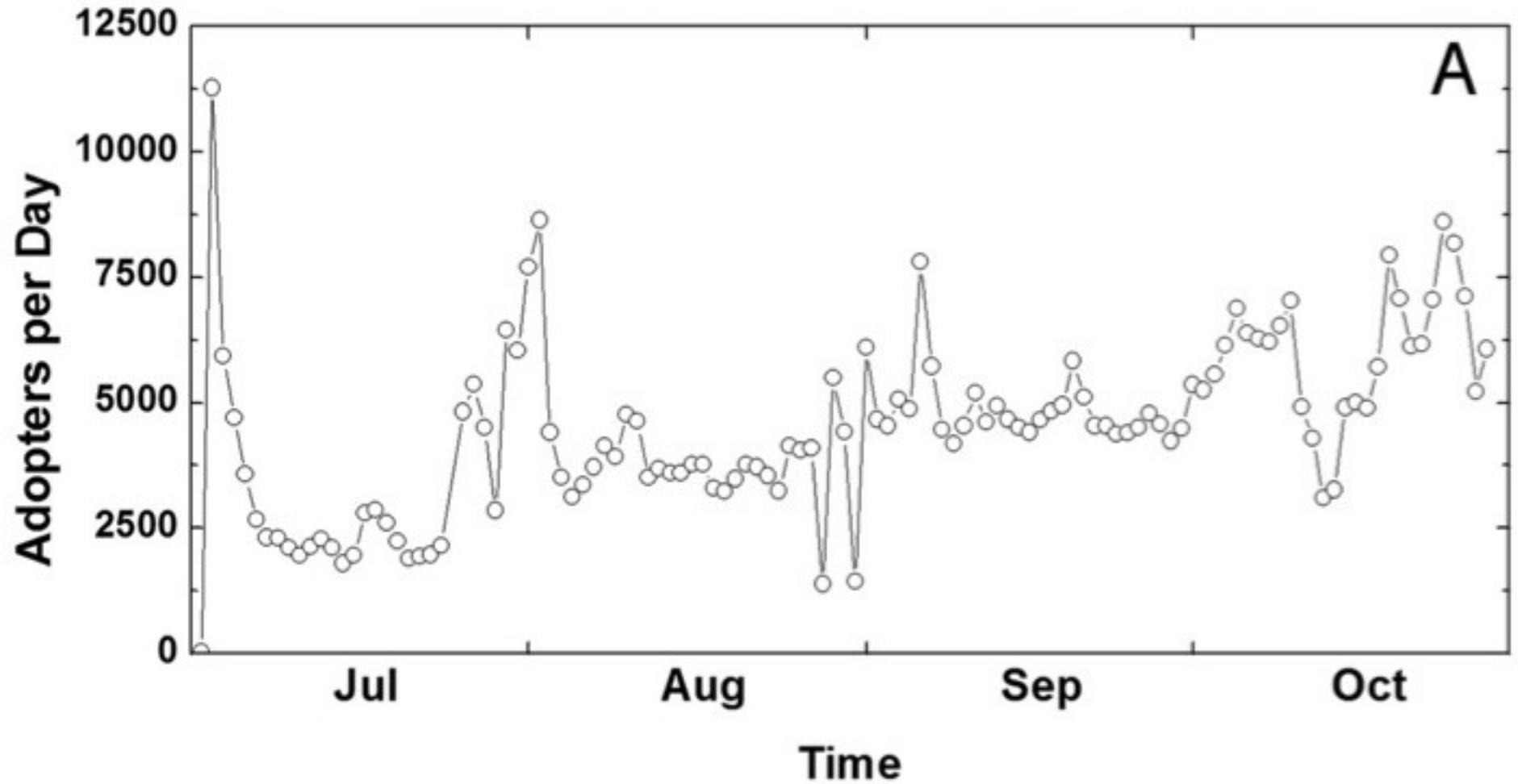
Propensity score matching (Aral et al.)

Yahoo! Go

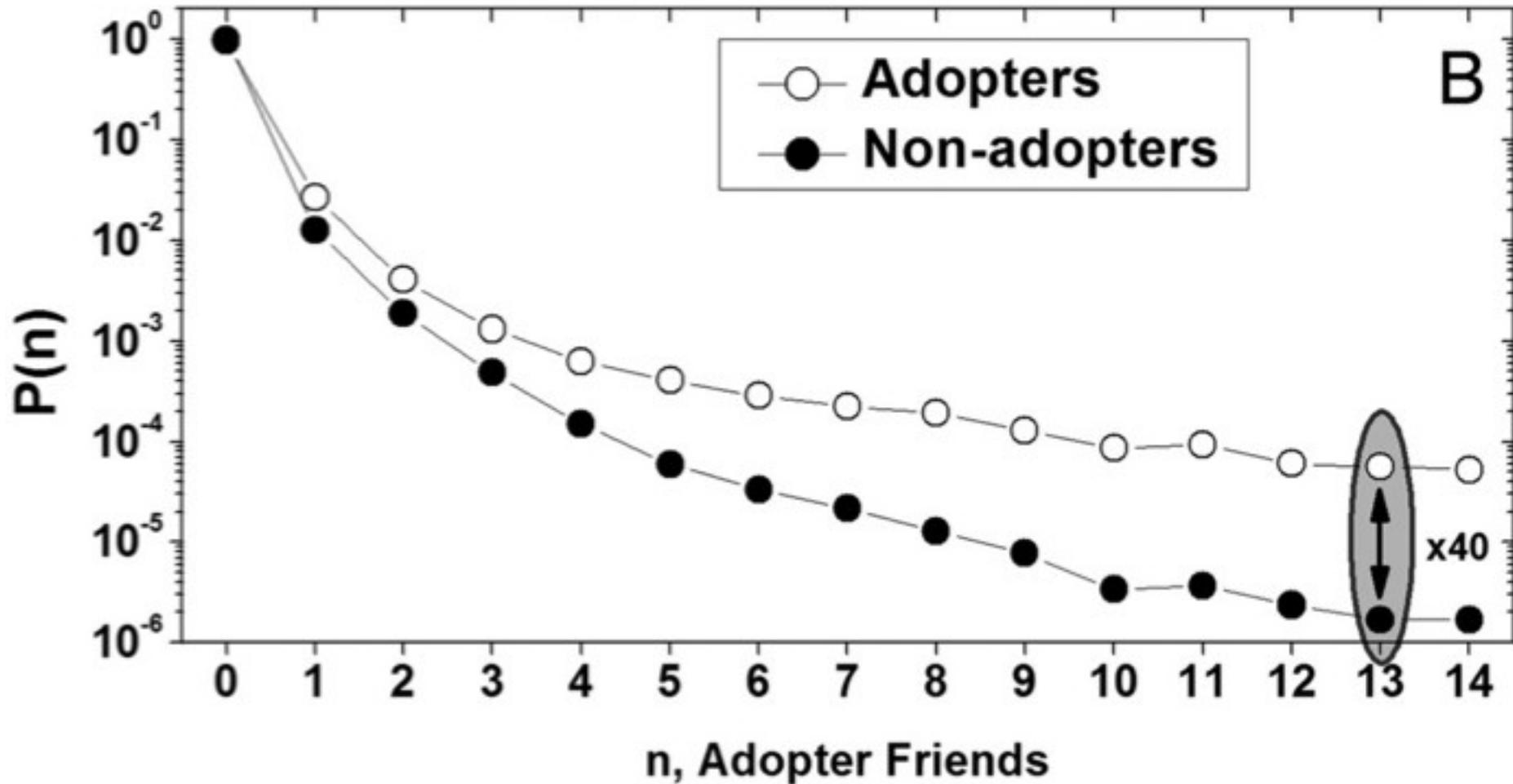


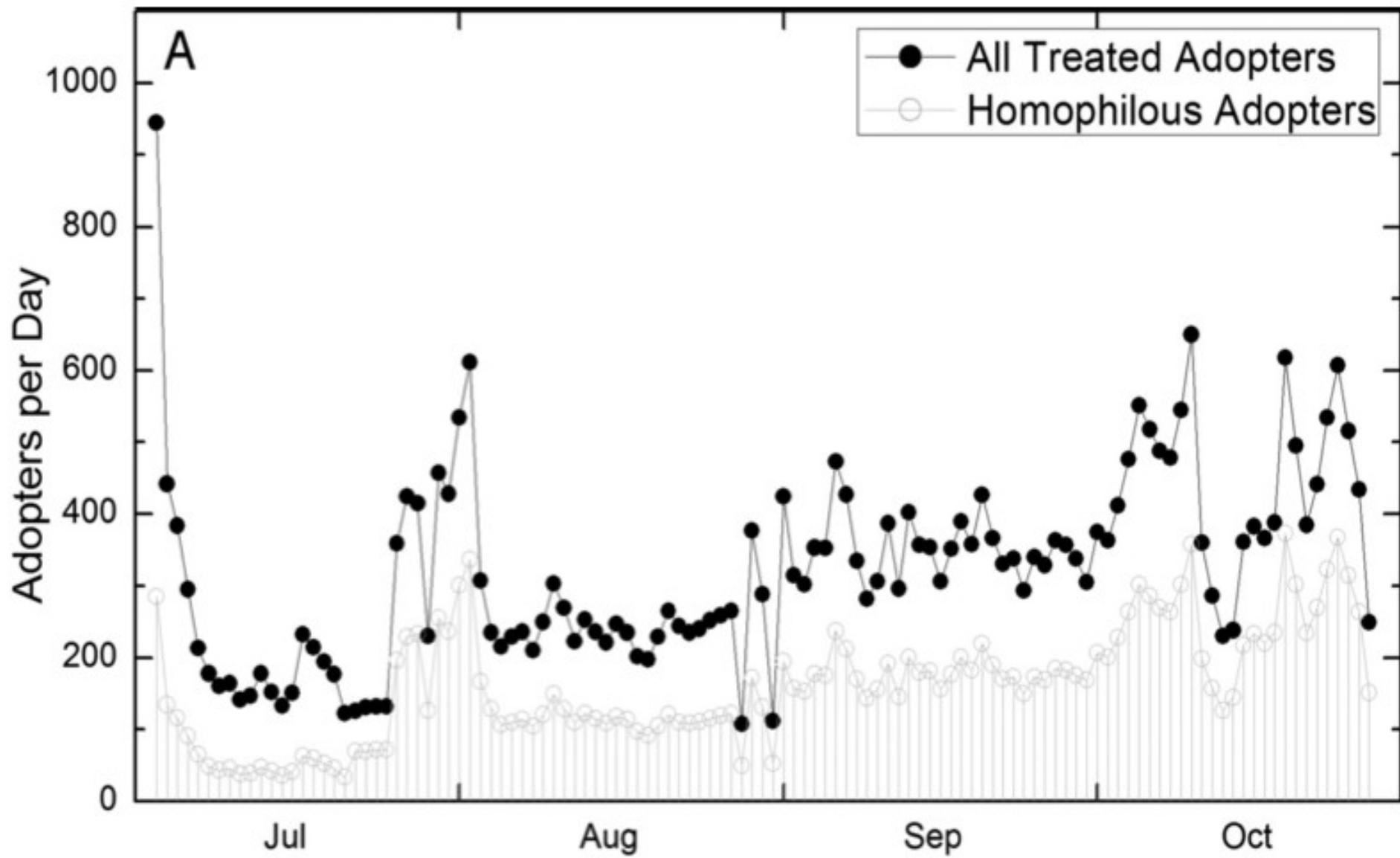
Dataset: 27.4 milioni di utenti Yahoo!

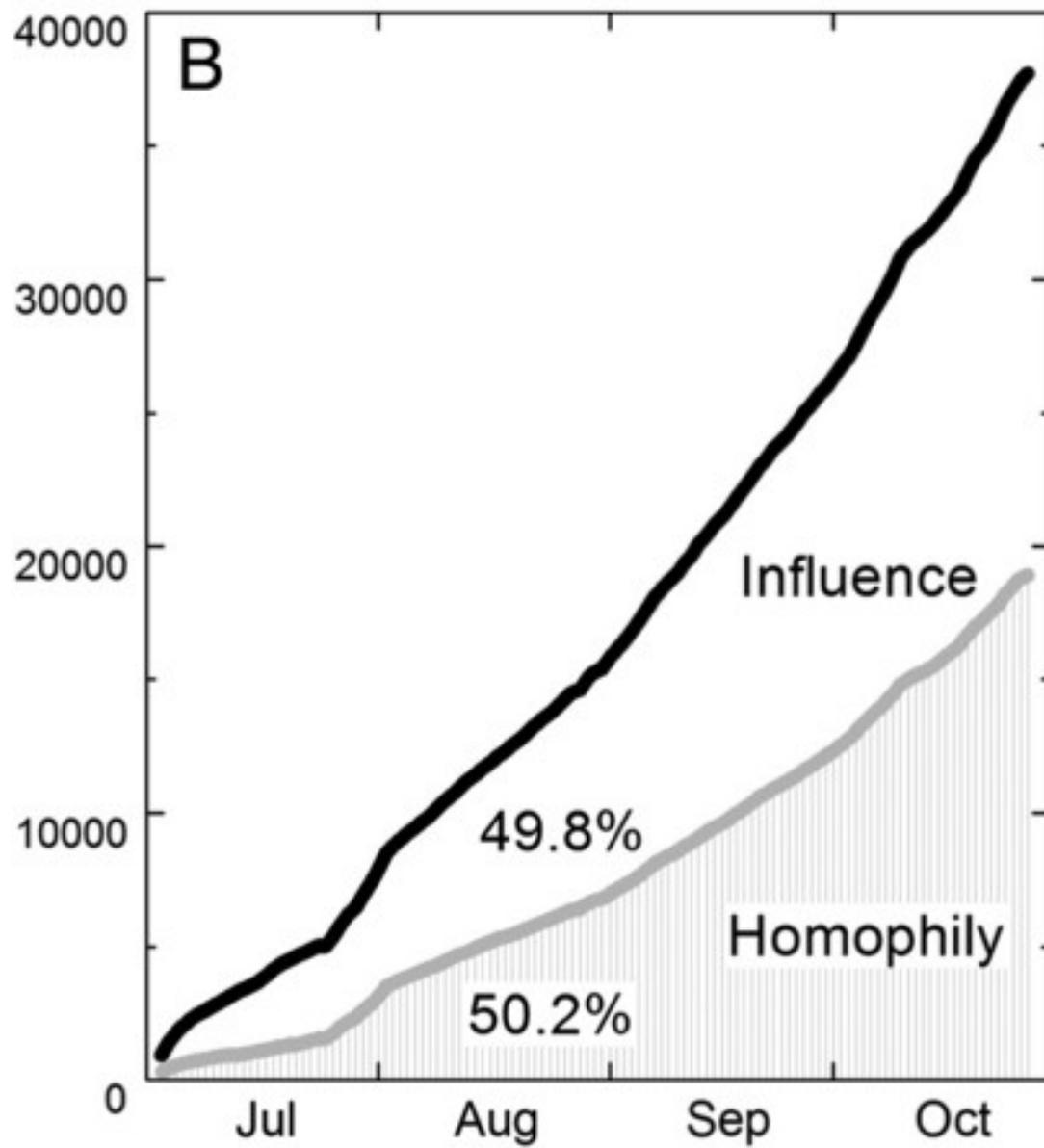
Adozioni per giorno



Probabilità di essere un Adopter







Conclusioni

- Omofilia (sociologia)
- Assortative mixing (misurare l'omofilia)
 - Reti sociali
 - Reti tecnologiche/biologiche
- Assortative mixing e robustezza reti
- Influenza dell'omofilia nella diffusione dei comportamenti

Riferimenti

- M. McPherson, L. Smith-Lovin, J.M. Cook. Birds of a Feather: Homophily in Social Networks
- M. E. J. Newman. Assortative Mixing in networks.
- M. E. J. Newman. Mixing Patterns in networks
- S. Aral, L. Muchnik, A. Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic network
- C. R. Shalizi, A. C. Thomas. Homophily and Contagion Are Generically Confounded in Observational Social Network