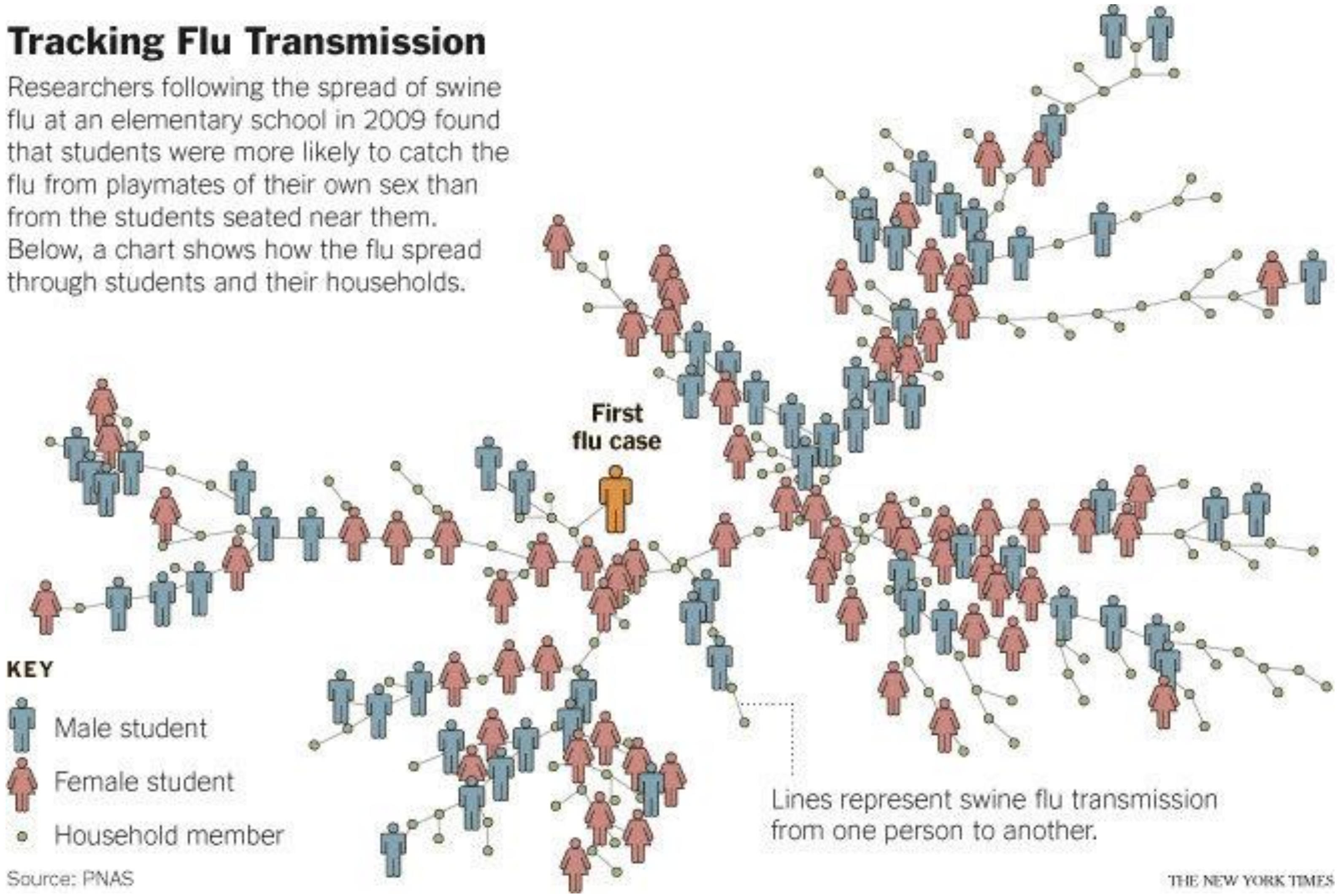
Social network analysis

Social contagion

Social distance trumps physical distance



“Ideas and products and message and behaviors spread just like viruses do”

(Malcolm Gladwell, The Tipping Point)

Social epidemics, like infectious diseases, are passed along by a handful of exceptional actors

Meme (“Memory” + “gene”) the mind “virus” (**Richard Dawkins**)

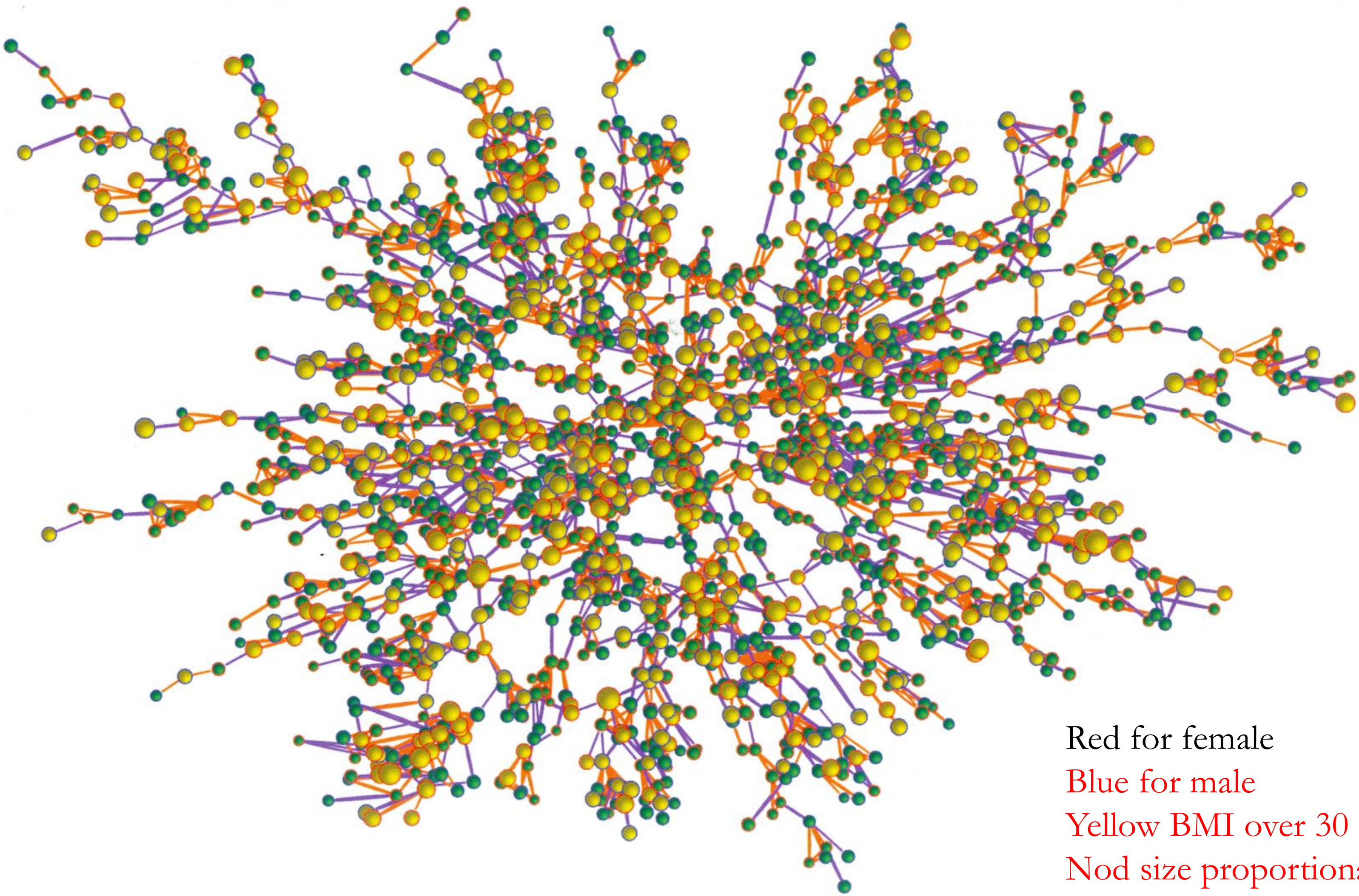
Any idea or behavior that can pass from one person to another by learning or imitation. Examples include thoughts, ideas, theories, gestures, practices, fashions, habits, songs, and dances

Obesity

- The average obese person was more likely to have friends, friends of friends, and friends of friends of friends who were obese than would be expected due to chance alone.
- The average nonobese person was, similarly, more likely to have nonobese contacts up to three degree of separation, beyond 3 degrees, the clustering stopped.

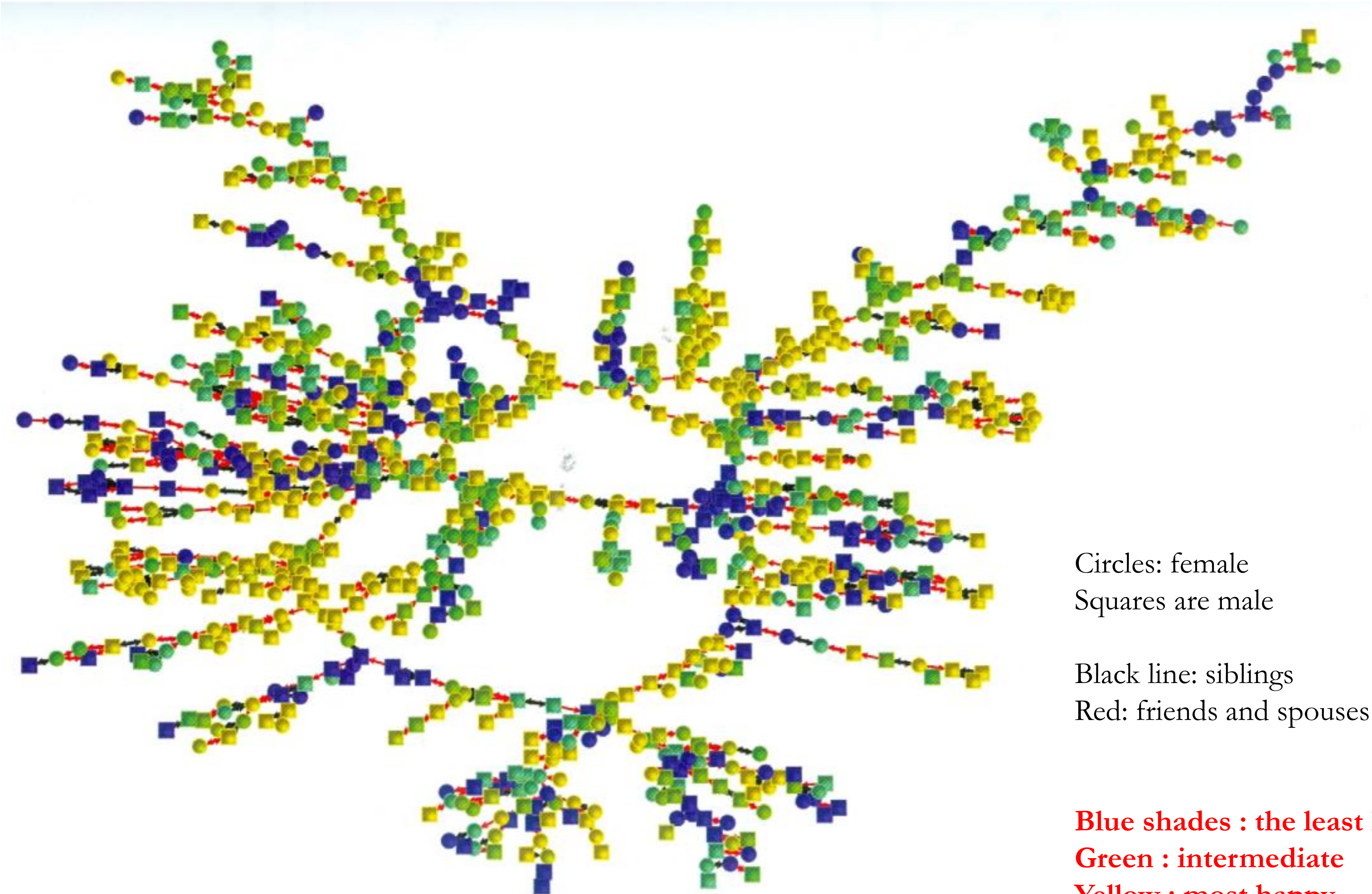
[Nicholas Christakis TED talk]

Obesity and social influences



Red for female
Blue for male
Yellow BMI over 30
Nod size proportional to BMI
Green BMI less than 30
Tie colors indicate relationships
Purple for friend or spouse,
orange for family

Happiness and social influence?

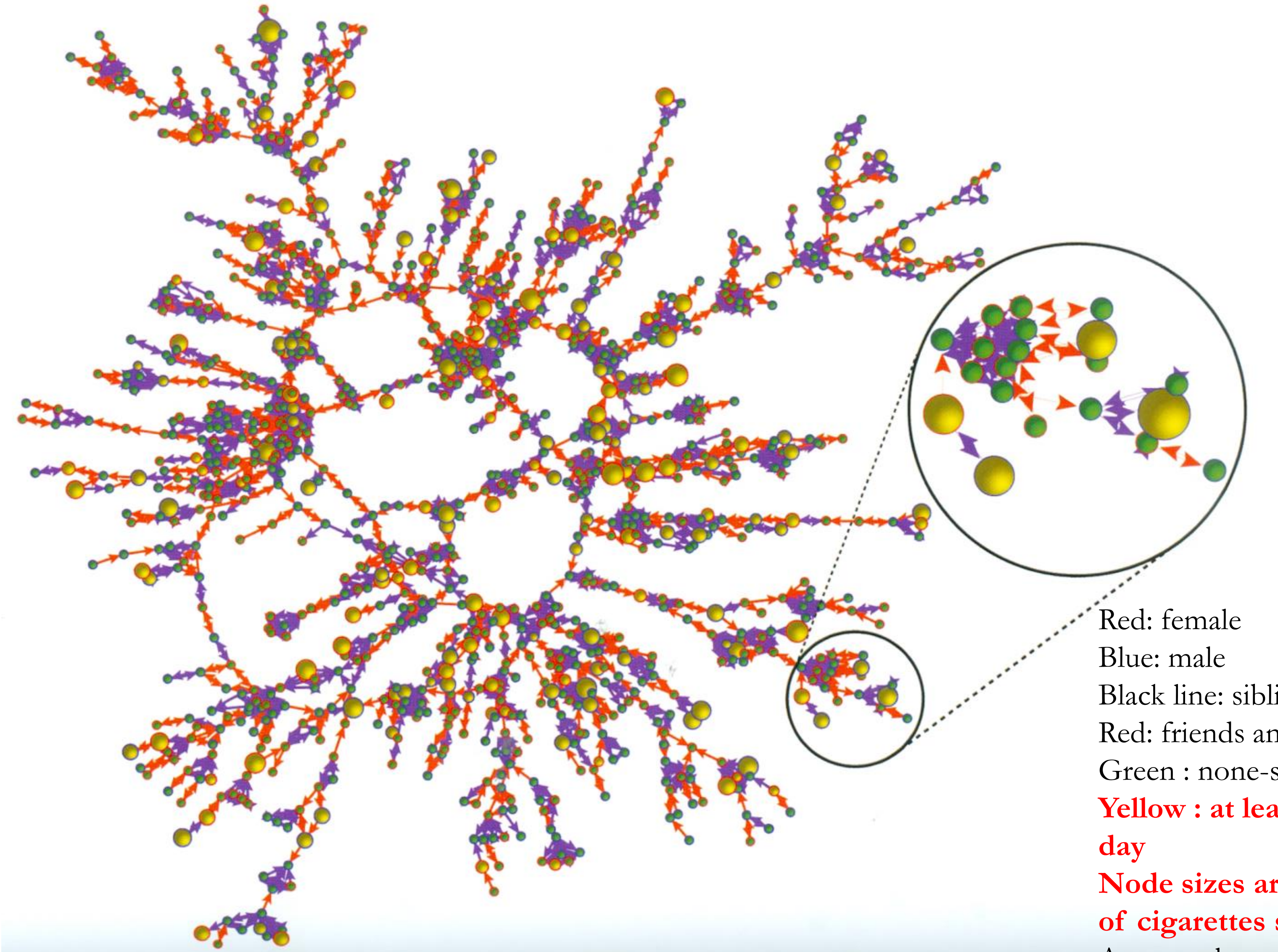


Circles: female
Squares are male

Black line: siblings
Red: friends and spouses

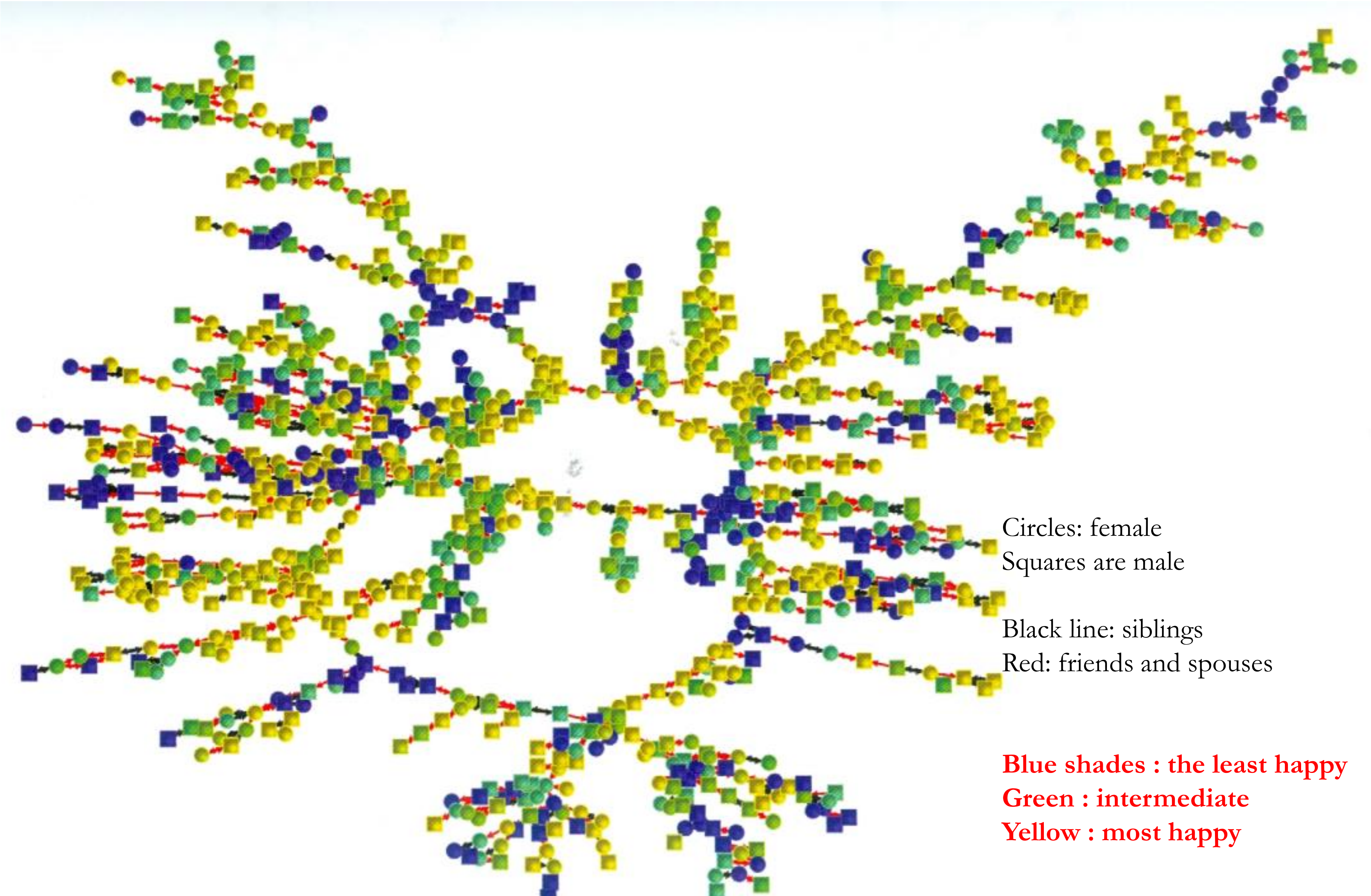
Blue shades : the least happy
Green : intermediate
Yellow : most happy

Smoking and social influence



Red: female
Blue: male
Black line: siblings
Red: friends and spouses
Green : none-smoker
Yellow : at least one cigarette a day
Node sizes are proportional to # of cigarettes smoked
Arrow colors orange : friends and spouse
Purple for family

Happiness and social influence?

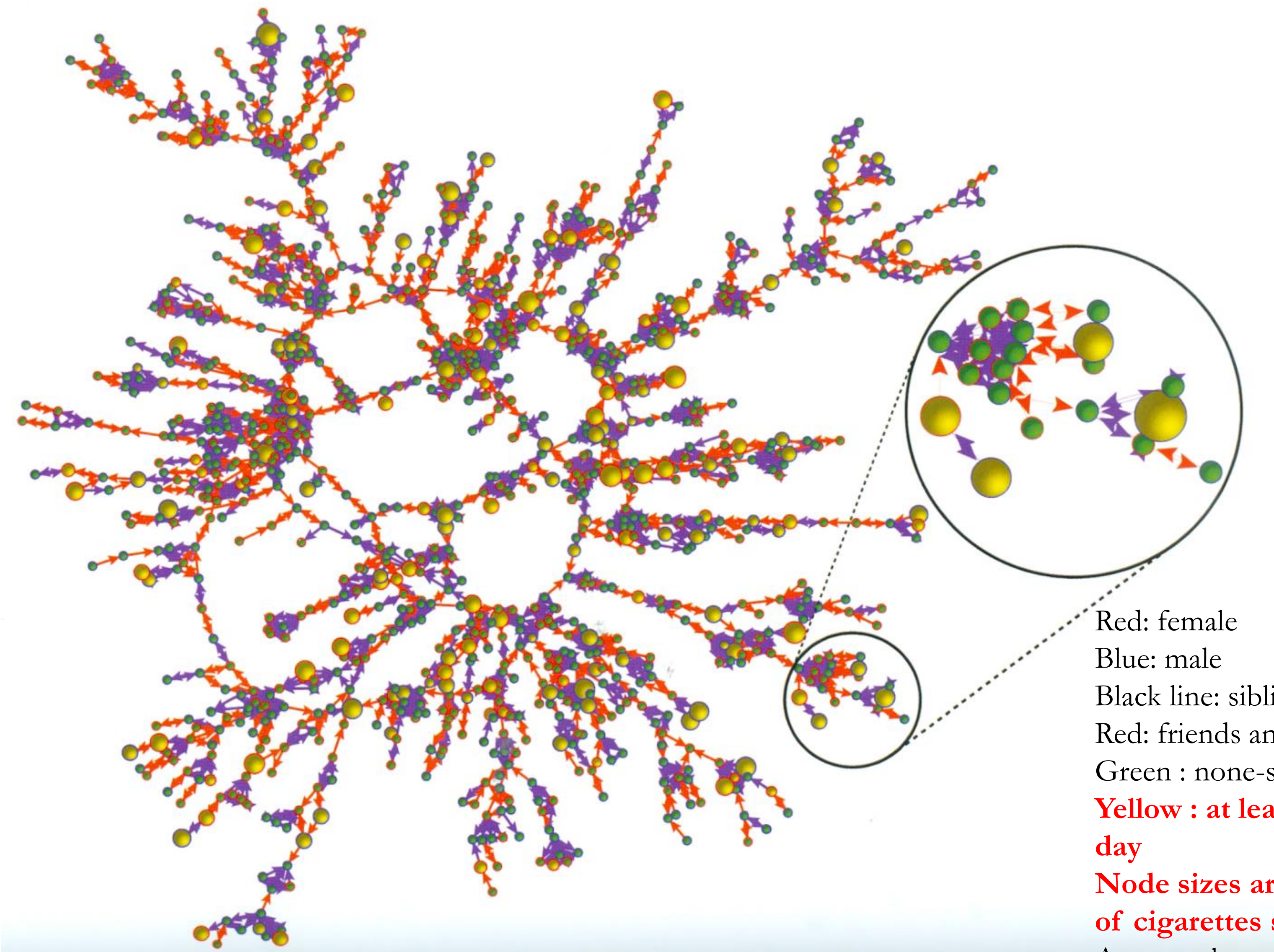


Circles: female
Squares are male

Black line: siblings
Red: friends and spouses

Blue shades : the least happy
Green : intermediate
Yellow : most happy

Smoking and social influence



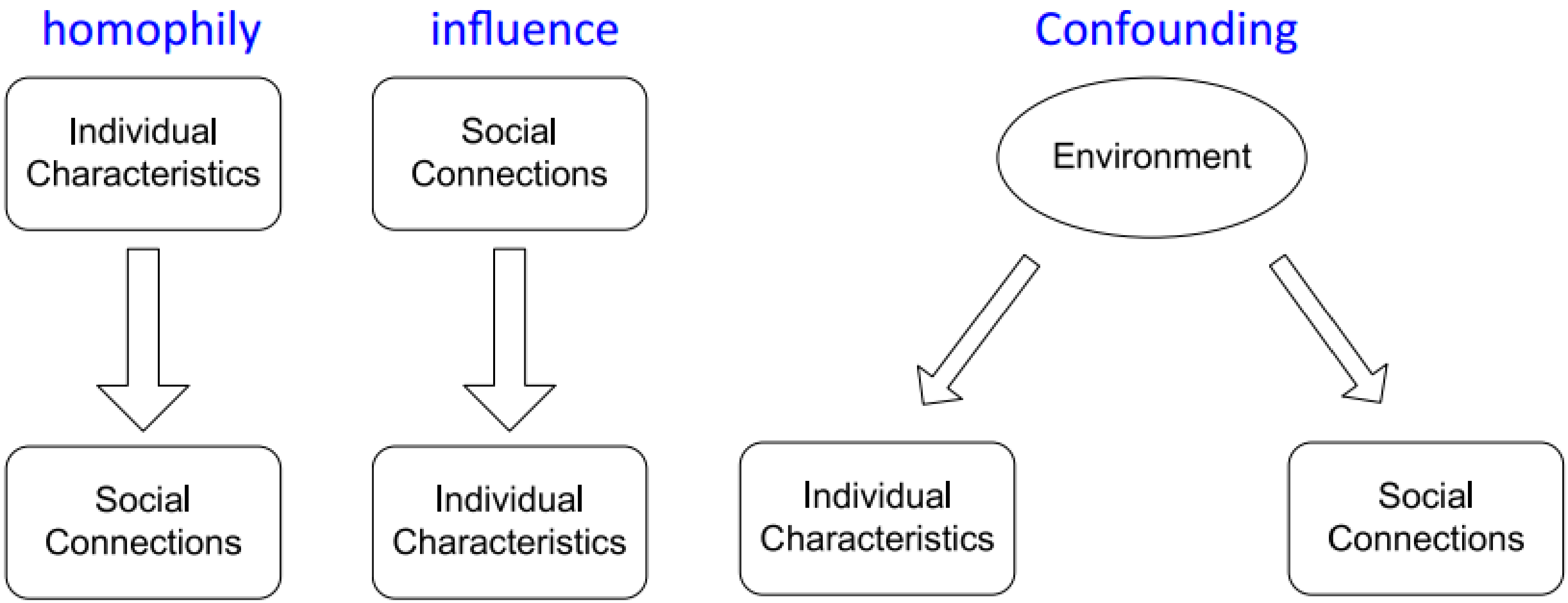
Red: female
Blue: male
Black line: siblings
Red: friends and spouses
Green : none-smoker
Yellow : at least one cigarette a day
Node sizes are proportional to # of cigarettes smoked
Arrow colors orange : friends and spouse
Purple for family

Some observations of life in the Network (Christakis and Fowler, 2009)

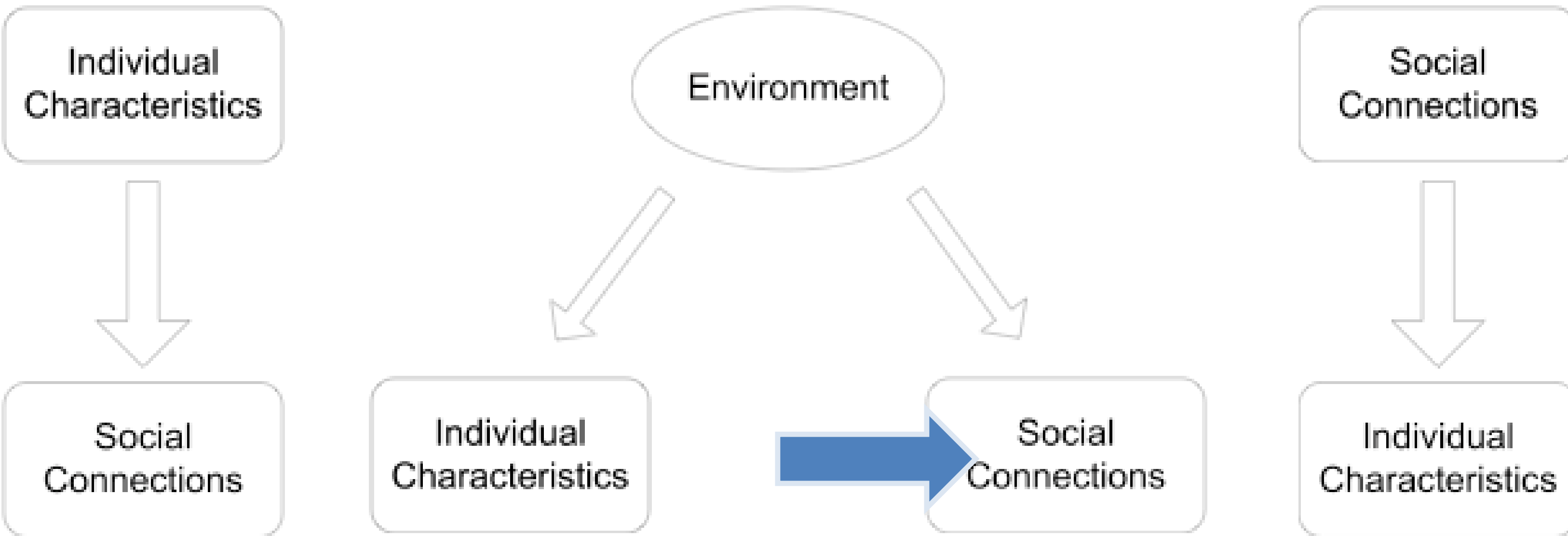
- We shape our network
  - Homophile or selection: “love of being alike”
- Our network shapes us
  - Social conformity, imitation, network effect
- Our friends affect us
- Our friends’ friends affect us
- The network work has a life of its own

Structure and agency

- Network analysis allows us to examine how the configuration of networks influences how individuals and groups, organizations, or systems function.


homophily
influence
Confounding
Individual Characteristics
Social Connections
Social Connections
Individual Characteristics
Environment
Individual Characteristics
Social Connections

homophily — Confounding — Social contagion/influence

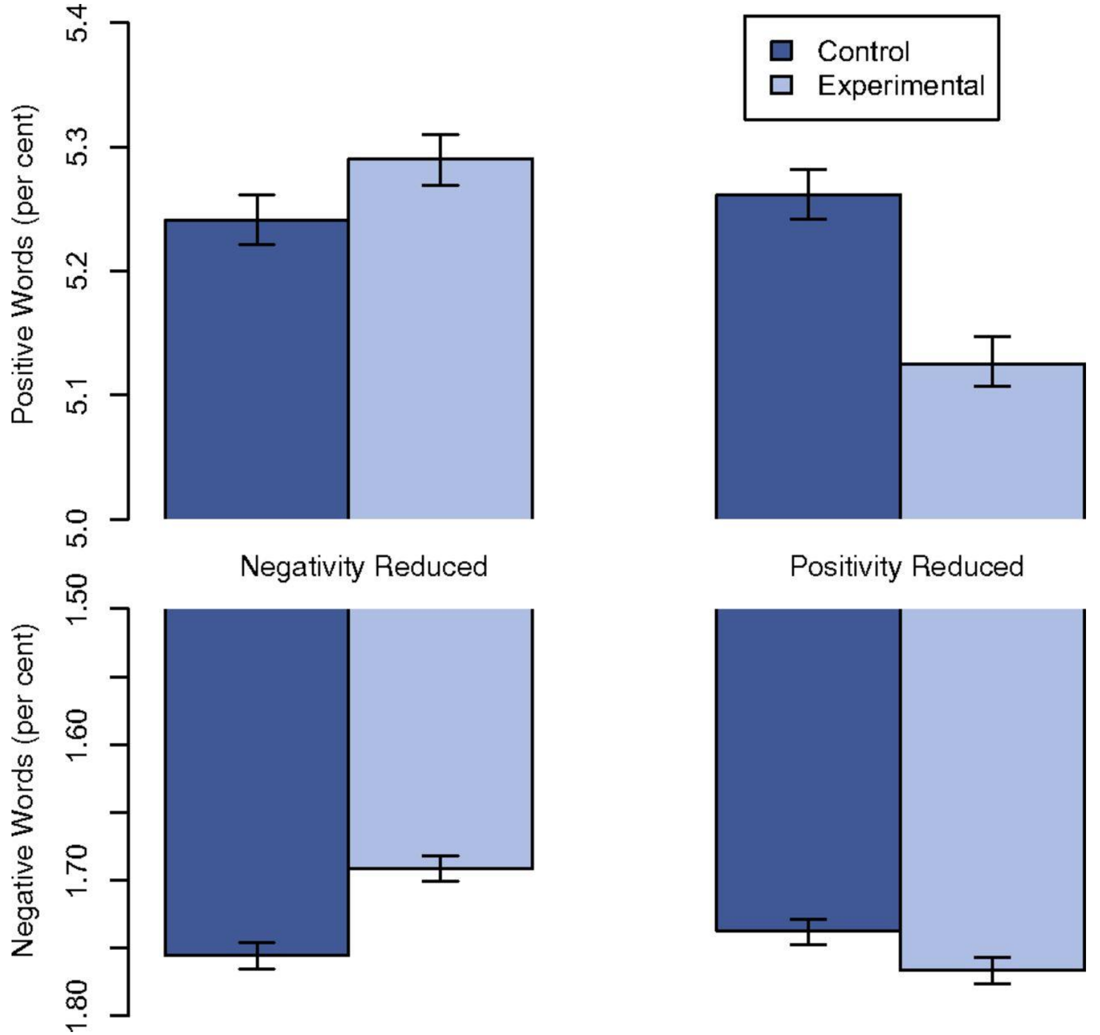


Individual characteristics
e.g. smoking, obesity
Reading profile ?

Experimental evidence of massive-scale emotional contagion through social networks (PNAS)

- Took place January 11-18, 2012, N= 689,003
- "One in which exposure to friends' positive emotional content in their News Feed was reduced, and one in which exposure to negative emotional content in their News Feed was reduced "
- And a "control condition, in which a similar proportion of posts in their News Feed were omitted entirely at random

Mean number of positive (Upper) and negative (Lower) emotion words (percent) generated people, by condition.



Kramer A D I et al. PNAS 2014;111:8788-8790

Random sample of non-interacting pairs

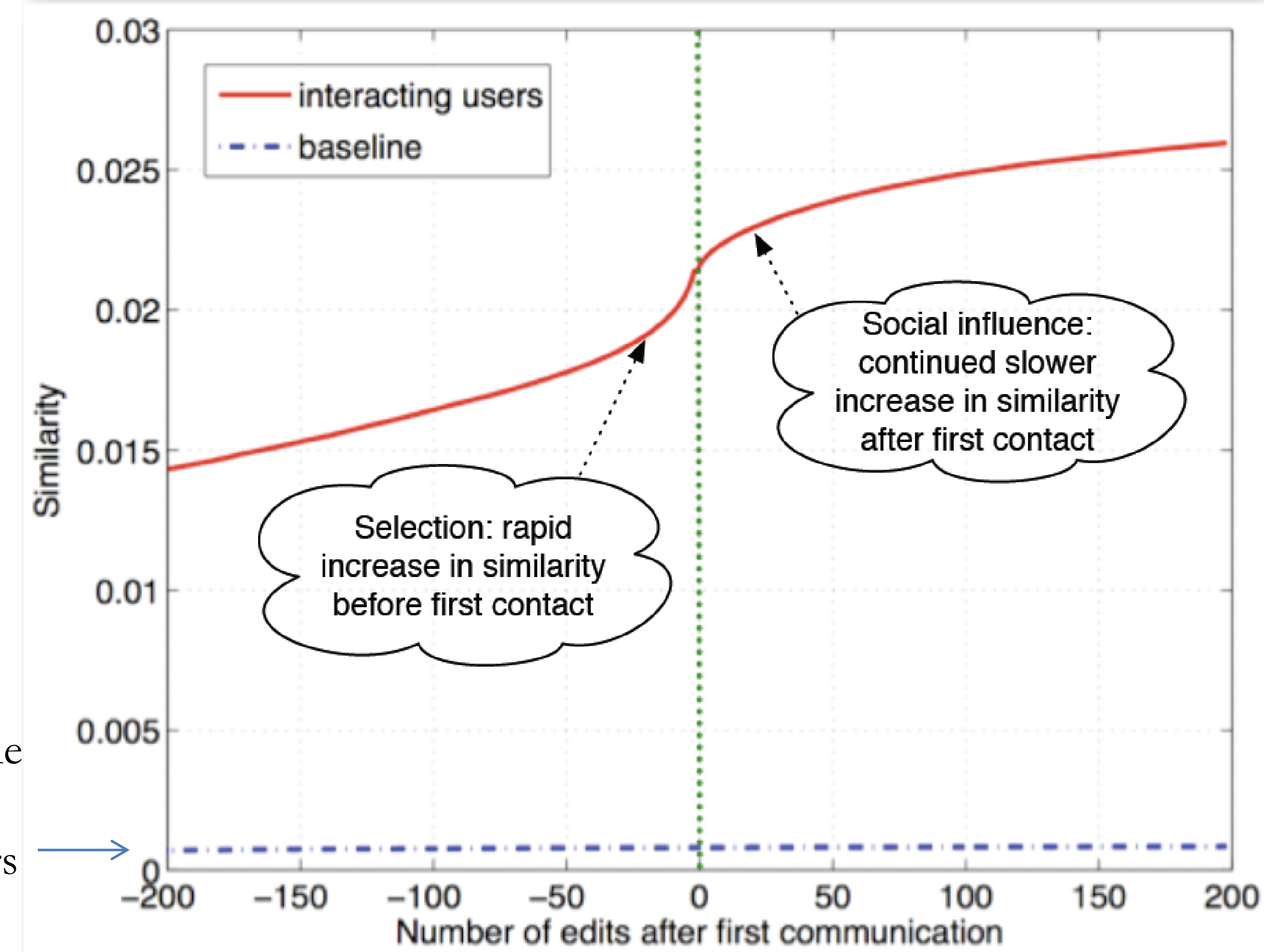


Figure 4.13: The average similarity of two editors on Wikipedia, relative to the time (0) at which they first communicated [122]. Time, on the x-axis, is measured in discrete units, where each unit corresponds to a single Wikipedia action taken by either of the two editors. The curve increases both before and after the first contact at time 0, indicating that both selection and social influence play a role; the increase in similarity is steepest just before time 0.

- Link creation and profile alignment in the aNobii social network

  Luca Maria Aiello, Alain Barrat, Ciro Cattuto,
  Giancarlo Ruffo and Rossano Schifanella
  IEEE SocialCom 2010

Outline

- Introduction of aNobii
- Static datasets analysis
  - Static properties of user's OSN and activity.
  - Geographic features.
  - Role of profile similarity
- Dynamic dataset analysis
  - How users create new relationship
    - Homophily vs. social influence
  - Network growth

aNobii


anobii
登錄 | 註冊帳號
Login
Together we find better books
書櫃
尋找
分享
書櫃
放上書櫃並展示您的藏書
尋找
藉由 aNobii 社群的協助，找尋您的下一本書籍
分享
評分並評論您的藏書，並在 Facebook 與 Twitter 上分享給大家
您現在正在閱讀哪本書？
搜尋
近期活動
哈寶寶在書架上新增了皇權禍國
14 分鐘前
const在書架上新增了視覺溝通的文法
14 分鐘前
Roxanne在書架上新增了從巴黎到巴塞隆納，慢慢走
14 分鐘前
Ariel4316 在書架上新增了談美感
14 分鐘前
voldemort在書架上新增了實用德語會話
15 分鐘前
青蛙王子在書架上新增了乘業日記
15 分鐘前
1樓書櫃在書架上新增了龍眠
17 分鐘前
hsinyun在書架上新增了一個人的第一次
17 分鐘前
aNobii
for iPhone
Enhanced Barcode Scanning
Access Expert Book Reviews
iPhone 4 Ready
aNobii 應用程式
iPhone
Android
掃描任一條碼來建立您的虛擬書櫃，並了解每個人對您藏書的看法
關於 | 翻譯 | 開發者 | Blog | 說明 | 條款 | 私隱 | 工作 | 聯絡
© 2012 aNobii
繁體中文

Social network for bookworms

- Data-driven analysis on anobii.com
- Profile features
  - Library and wish list
  - Groups
  - Tags
- Social network
  - Directed
  - Friendship + neighborhood

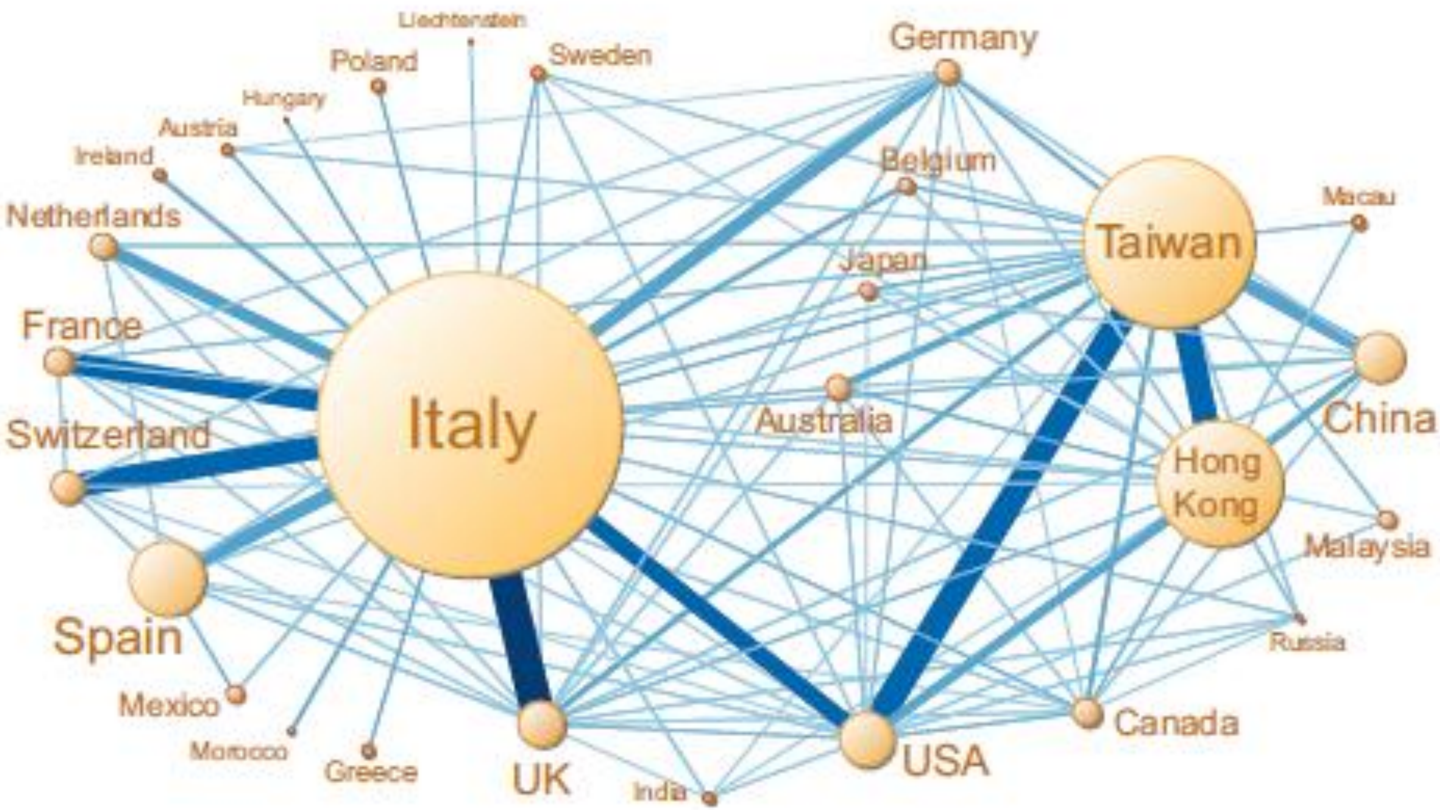
4th snapshot	Friendship	Neighborhood	Union
Nodes	74,908	54,590	86,800
Links	268,655	429,482	697,910

- 6 snapshots, 15 days apart
- *Full* giant connected component

Outline

- Introduction of aNobii
- Static datasets analysis
  - Static properties of user's OSN and activity.
  - Geographic features.
  - Role of profile similarity
- Dynamic dataset analysis
  - How users create new relationship
    - Homophily vs. social influence
  - Network growth

Geographical analysis



Dataset analysis

- Dataset crawling
  - From a random seed
  - Neighborhood network and friendship network
  - 6 snapshot starting from 11/09/2009
- Static properties
  - K-out degrees
  - SCC, WCC
  - Average shortest path length

Static Dataset analysis

	Friendship	Neighborhood	Union
Nodes	74,908	54,590	86,800
Links	268,655	429,482	697,910
Reciprocation	0.71	0.45	0.57
$< k_{out} >$	3.6	7.9	8.0
WCC size	68,624	54,246	86,800
SCC size	46,253	29,110	62,195
Density	$4.8 \cdot 10^{-5}$	$1.4 \cdot 10^{-4}$	$9.3 \cdot 10^{-5}$
Average SPL	7.3	4.7	5.3
Diameter	25	15	20
Degree centr.	0.0082	0.12	0.079

TABLE I
FRIENDSHIP, NEIGHBORHOOD AND FULL SOCIAL NETWORK STATISTICS (SPL=SHORTEST PATH LENGTH; WCC=WEAKLY CONNECTED COMPONENT; SCC=STRONGLY CONNECTED COMPONENT).

Static Dataset analysis (cont.)

- Diameter: The longest shortest path between two nodes in a network.
- Neighborhood network is denser and with a higher degree centralization
- Statistics reflects that the interest in other users' readings tends to concentrate toward a core.

Static Dataset analysis: power law distribution

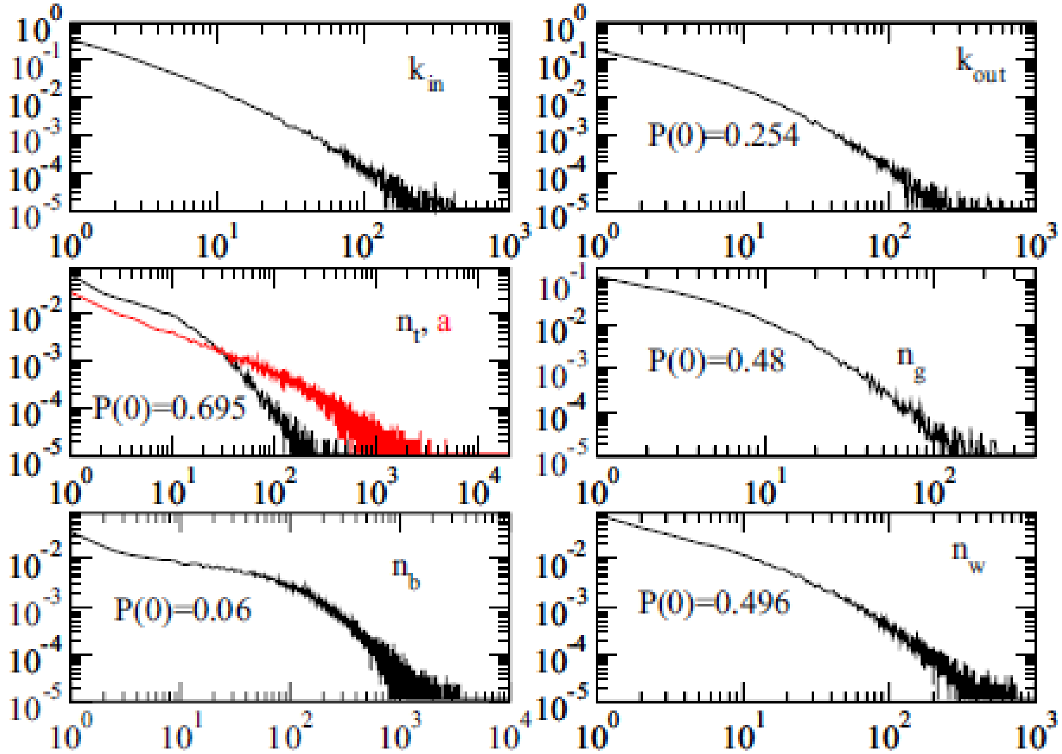
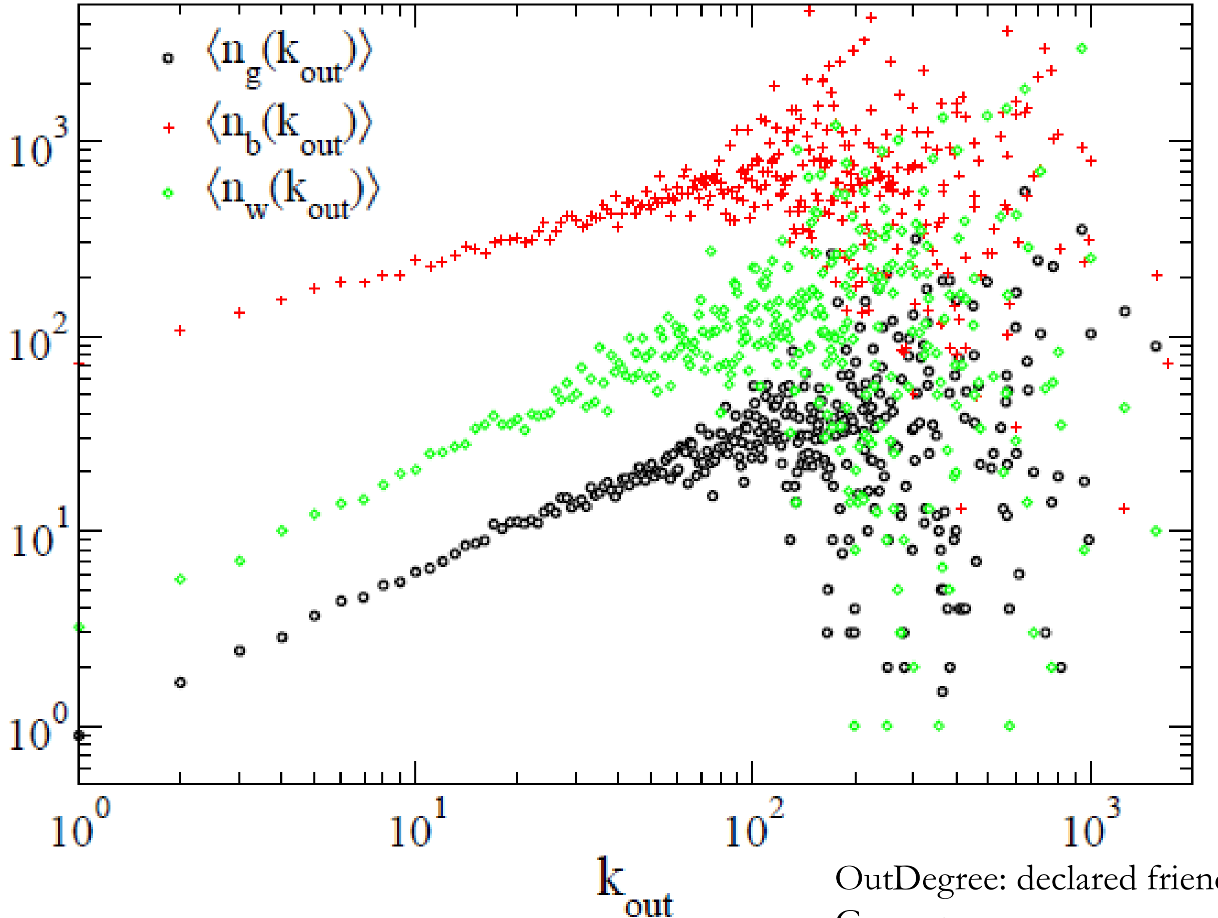


Fig. 1. Distributions of the measures of activity of aNobii users: in-degree k_{in} and out-degree k_{out} in the social network, number of distinct tags n_t and total tagging activity a (total number of tags in a user's page), number of group memberships n_g, number of books in a user library n_b and in a user wishlist n_w.

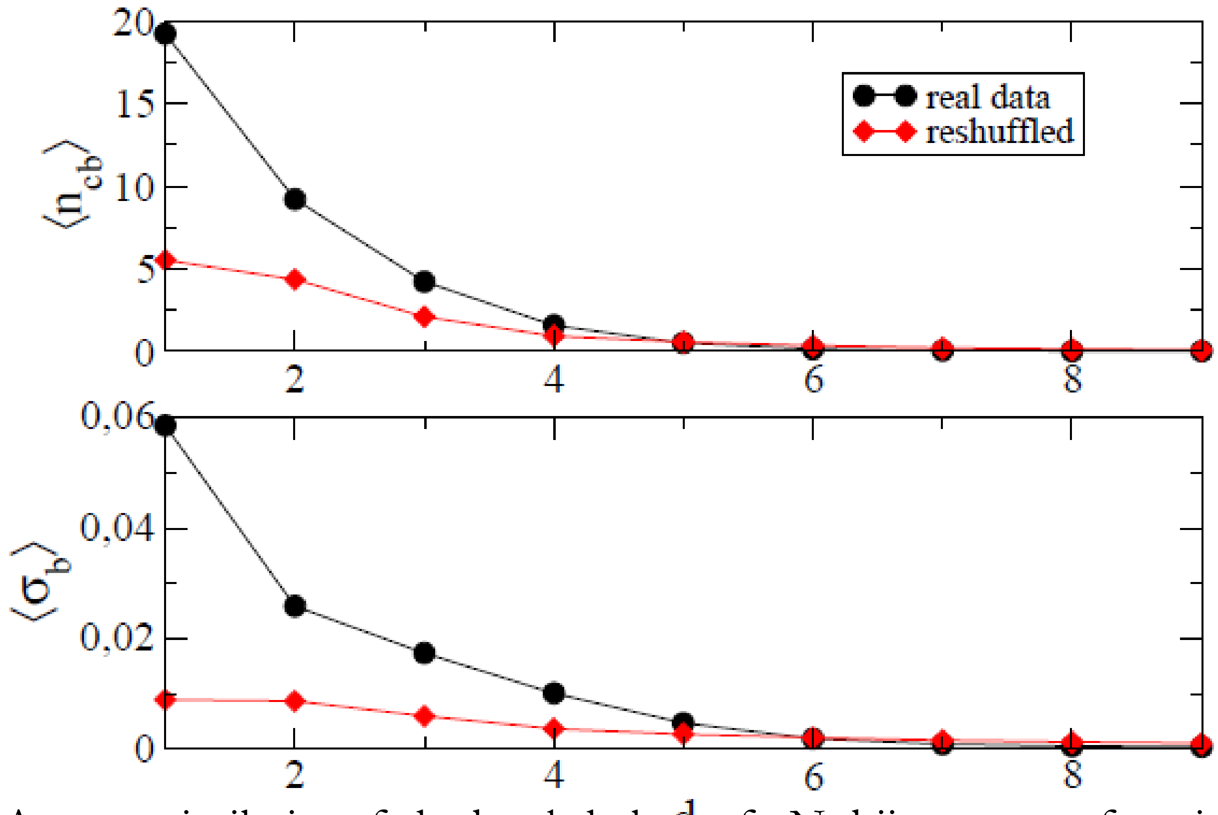
Correlations between one's outdegrees and other activities



OutDegree: declared friends and neighbors
G:groups
B:books in shelves
W:books in wish list

Profile similarity and social distance

Does similarity between user profiles depend on the social distance?



Average similarity of the bookshelves of aNobii users as a function of their Distance in the social network.

Up: Average number of shared books

Bottom: Cosine similarity of bookshelves

Reshuffled: randomization

$$\sigma_b(u,v) = \frac{\sum_b \delta_u(b)\delta_v(b)}{\sqrt{n_b(u)n_b(v)}}$$

Outline

- Introduction of aNobii
- Static datasets analysis
  - Static properties of user's OSN and activity.
  - Geographic features.
  - Role of profile similarity
- Dynamic dataset analysis
  - How users create new relationship
    - Homophily vs. social influence
  - Network growth

Dynamical analysis

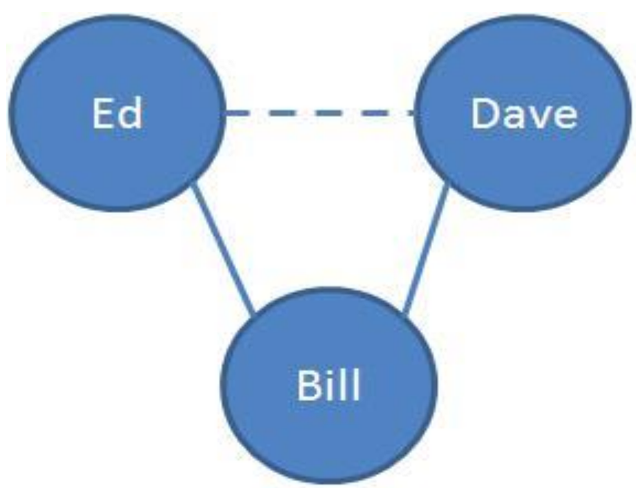
- Six snapshots in two and half months
- Dynamic analysis affords the study of
  - how people establish new relationship
  - network growth trend
  - new users’ behavior

Phenomena studied

- Triad closure
- Preferential attachment
- Homophily
- Social contagion

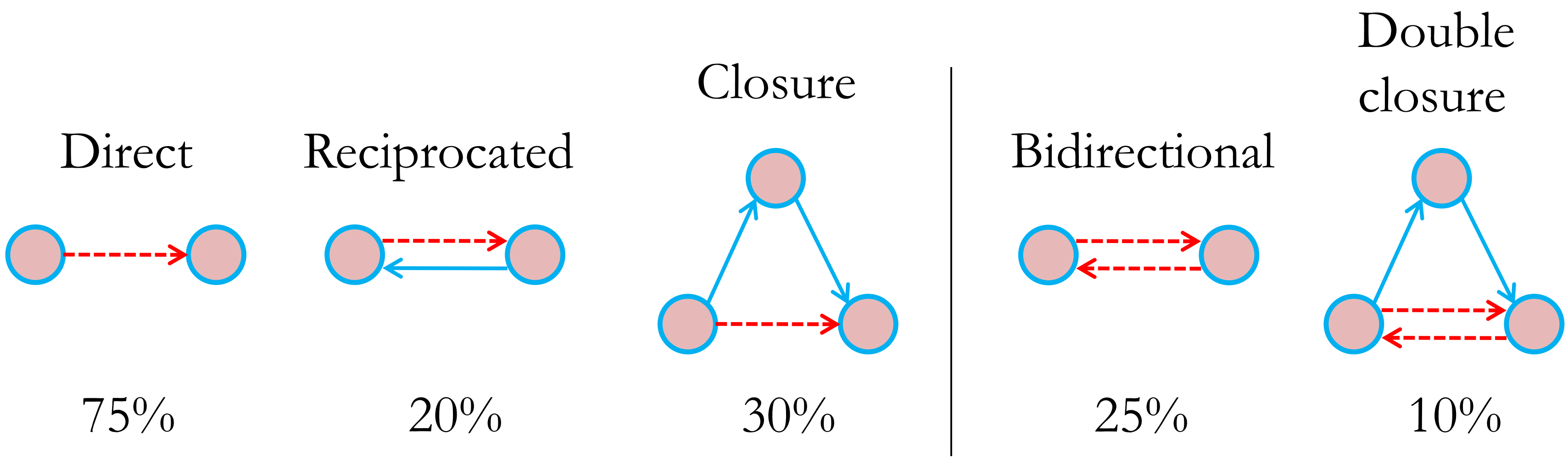
Triadic closure

- If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in future.



Triadic closure

- Classification of new links at time t+1 between nodes already present at time t (t ∈ {1,…,5})



- Reciprocation is strong
- Users tend to choose “friends of their friends” as new friends

Dynamical analysis: closure

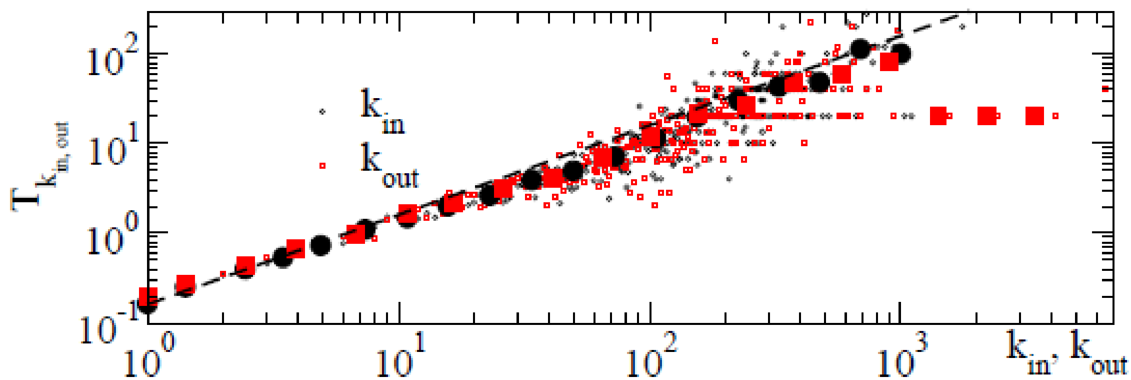
	$1 \to 2$	$2 \to 3$	$3 \to 4$	$4 \to 5$	$5 \to 6$
New nodes	2241	2121	1911	3214	3567
Removed nodes	239	222	230	220	684
New edges	19472	18324	17618	24805	26883
Died edges	642	763	713	782	700
$u \to v$	5409	4942	5259	6546	6357
Reciprocated	1016	1155	1285	1526	1688
$u \leftrightarrow v$	1809	1597	1604	1924	2235
Simple closure	2070	1976	2143	2497	2382
Double closure	955	904	877	1027	1141

TABLE III

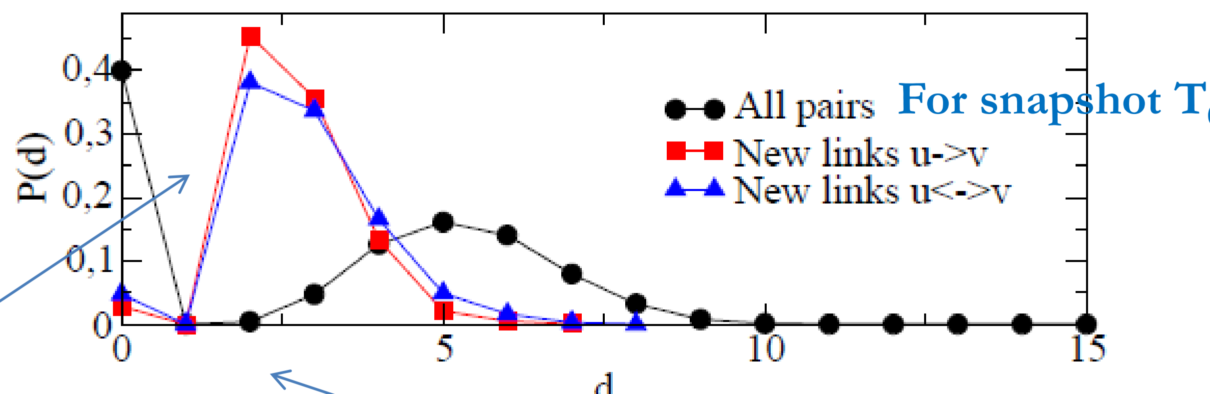
EVOLUTION OF SOME QUANTITIES FROM ONE SNAPSHOT TO THE NEXT.

Linkage bias: preferential attachment and homophily

Preferential attachment



Distance between newly created links between T4 and T5



Evidence for "selection" or "homophily"

Much shorter distance among linked pairs

Peaked at 2 degree,

Preferential attachment

- that the more connected a node is, the more likely it is to receive new links. Nodes with higher degree have stronger ability to grab links added to the network
- New nodes are added to the network one at a time. Each new node is connected to m existing nodes with a probability that is proportional to the number of links that the existing nodes already have.
- Formally, the probability p_i that the new node is connected to node i is

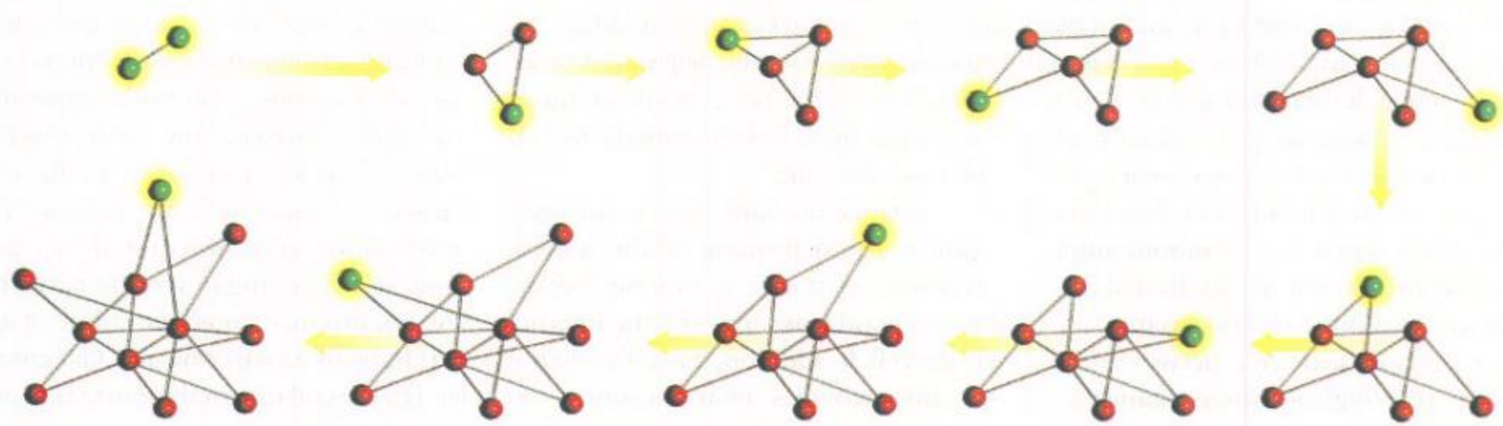
$$p_i = \frac{k_i}{\sum_j k_j},$$

Preferential attachment

- Network dynamics for evolving self-organized networks

BIRTH OF A SCALE-FREE NETWORK

A SCALE-FREE NETWORK grows incrementally from two to 11 nodes in this example. When deciding where to establish a link, a new node (*green*) prefers to attach to an existing node (*red*) that already has many other connections. These two basic mechanisms—growth and preferential attachment—will eventually lead to the system's being dominated by hubs, nodes having an enormous number of links.



Rich gets richer

Similarity → link creation

		$\langle n_{cb} \rangle$	σ_b	$\langle n_{cg} \rangle$	σ_g
At T_0	$d_{uv} = 2$	9.5	0.02	1.12	0.05
From T_0 To T_1	$u \longrightarrow v$	12.9	0.04	1.10	0.08
	$u \leftrightarrow v$	18.5	0.04	1.67	0.11
	Closure	18.2	0.04	1.81	0.10
	Dbl closure	23.4	0.05	1.20	0.12

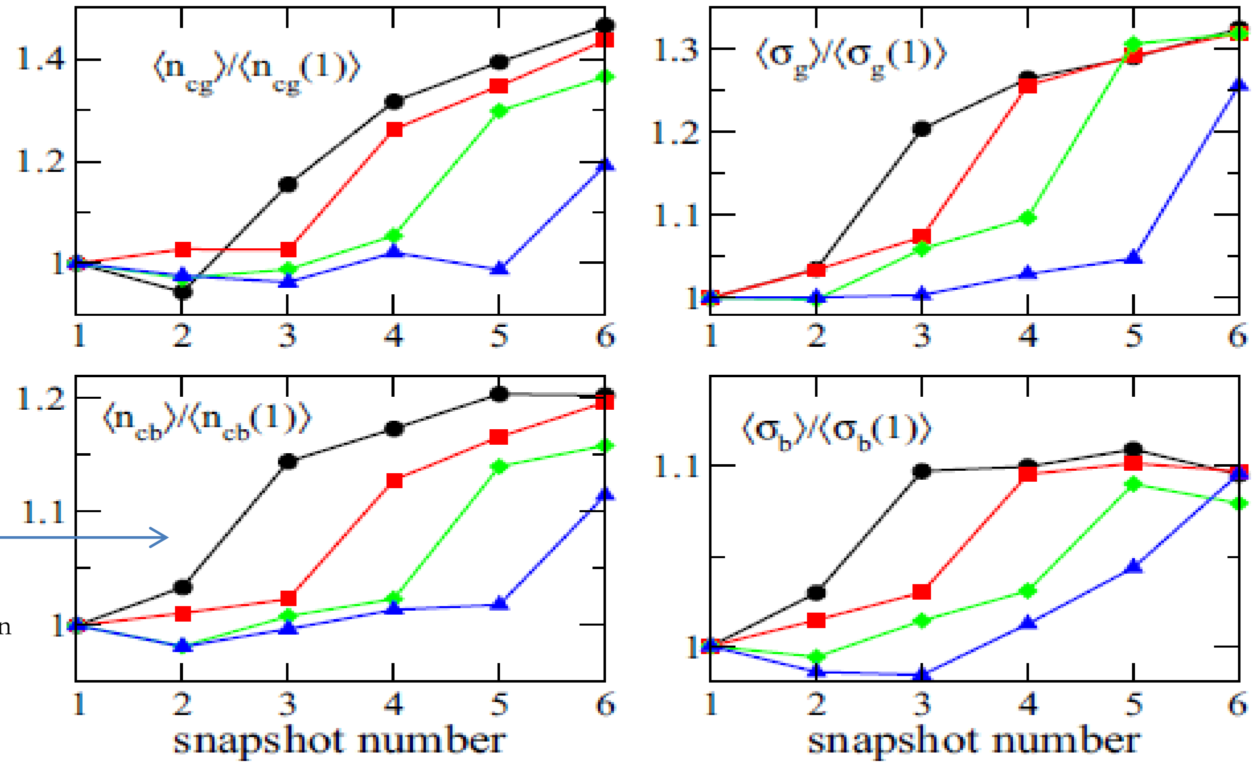
Average similarity of pairs forming new links between t_0 and t_0+1 ($t_0=4$), compared with average similarity of all the pairs at distance 2 at time t_0

Pairs that are going to get connected show a substantially higher similarity

Similarity of users partly drives the creation of new links

- “new links connect users who were already close, very often neighbors of neighbors; moreover, these users had more similar profiles than the average pairs of user at distance 2 ”

Link creation leads to similarity



Links formed between T2 and T3

Fig. 10. Evolution of the average similarity of users' profiles, as measured by the numbers of common books or groups, and by the cosine similarities, from the first to the last snapshot, $\boxed{\text{for links created between } t_0 \text{ and } t_0+1}$, for $t_0 = 2$ (black circles),3 (red squares), 4 (green diamonds), 5 (blue triangles), normalized by the average similarity in the first snapshot. Similarities are rather stationary before t_0, and clear jumps are observed between t_0 and $t_0 + 1$.

Social contagion

- “Before the creation of the links, the similarity is stationary; a large jump is observed when the links are created, and the similarity continues to grow, albeit at a slower page, after the link formation”
- “after users create links, they take inspiration from their new neighbors for new books to read and new group to join, and as a consequence align their profiles”

Aardvark and social search

- "a site which routes questions through an instant messaging bot to an appropriate user in one's extended network, comprised of friends-of-friends and strangers… prioritizes friends-of-friends for responses… When asked of the answers for feedback about the answer the user received**, 76% of the answers from within the user's extended network were rated 'good', and only 68% of answers from outside the users network were rated 'good'."**

Anyone flown Delta? Good? Bad?

March 13 at 9:17pm · Like · Comment

Pretty sure I've flown delta domestic in the US, from memory it was fine.

March 13 at 9:51pm · Like

There's about a hundred-dollar difference between them and United for three flights, two intl + one cross-country domestic...

March 13 at 9:57pm · Like

Delta is a thousand times better than united

March 13 at 9:58pm · Like

I agree, united are rubbish!

March 13 at 10:00pm · Like

Ok sweet, thanks guys! For Syd- USA, yeah?

March 13 at 10:05pm · Like

they're pretty much identical for domestic but i don't know anything about their international flights. did fly domestic or int'l? will we ever know? but yeah, generally speaking they're equivalent airlines.

also: consider the amazing AIR NEW ZEALAND. also the most important thing imo is to go to seatguru.com while you're booking your ticket so you can pick a good seat

March 13 at 10:07pm · Like

Yep, I'm all over seatguru (Expedia has it