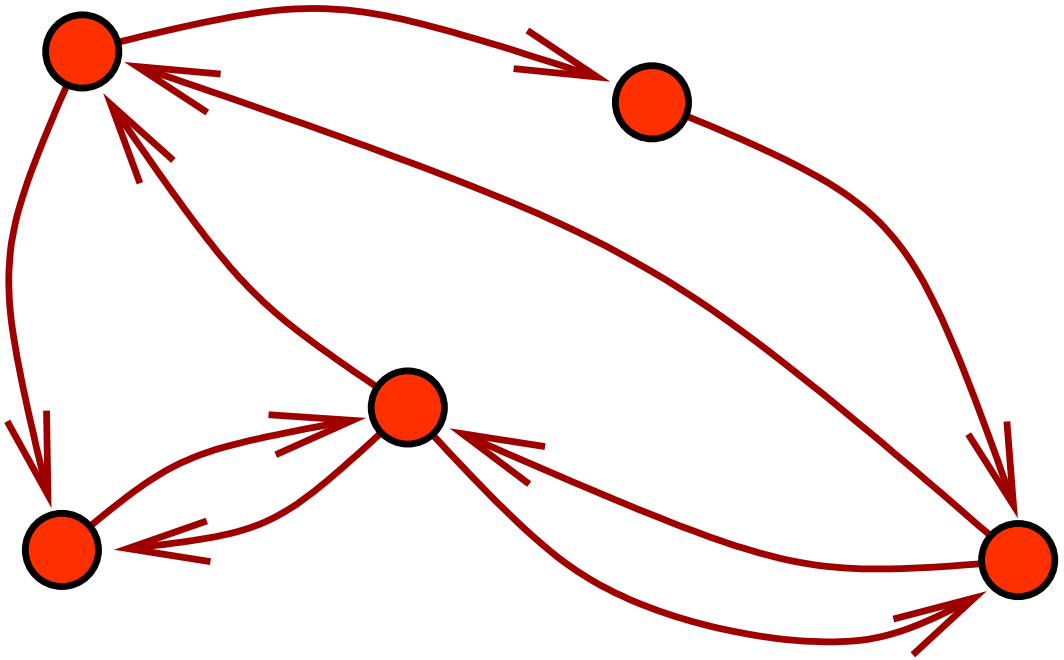
Social networks for physicists . . . and . . . the structure of an Internet dating community

January X, 2003

Petter Holme
Department of Physics,
Umeå University

Fredrik Liljeros &
Christofer Edling
Swedish Institute for
Infectious Disease Control
Department of Sociology,
Stockholm University

NETWORKS



A *network* is any system that can be modeled as a *graph*—a set of *vertices* and *arcs* (or *edges*).

The in/out-degree of a vertex is the number of arcs going in/out.

SOCIAL NETWORKS

- from group psychology . . .
- . . . to graph theory
- most intangible of networks
- epidemiology
- anthropology
- psychology
- defense

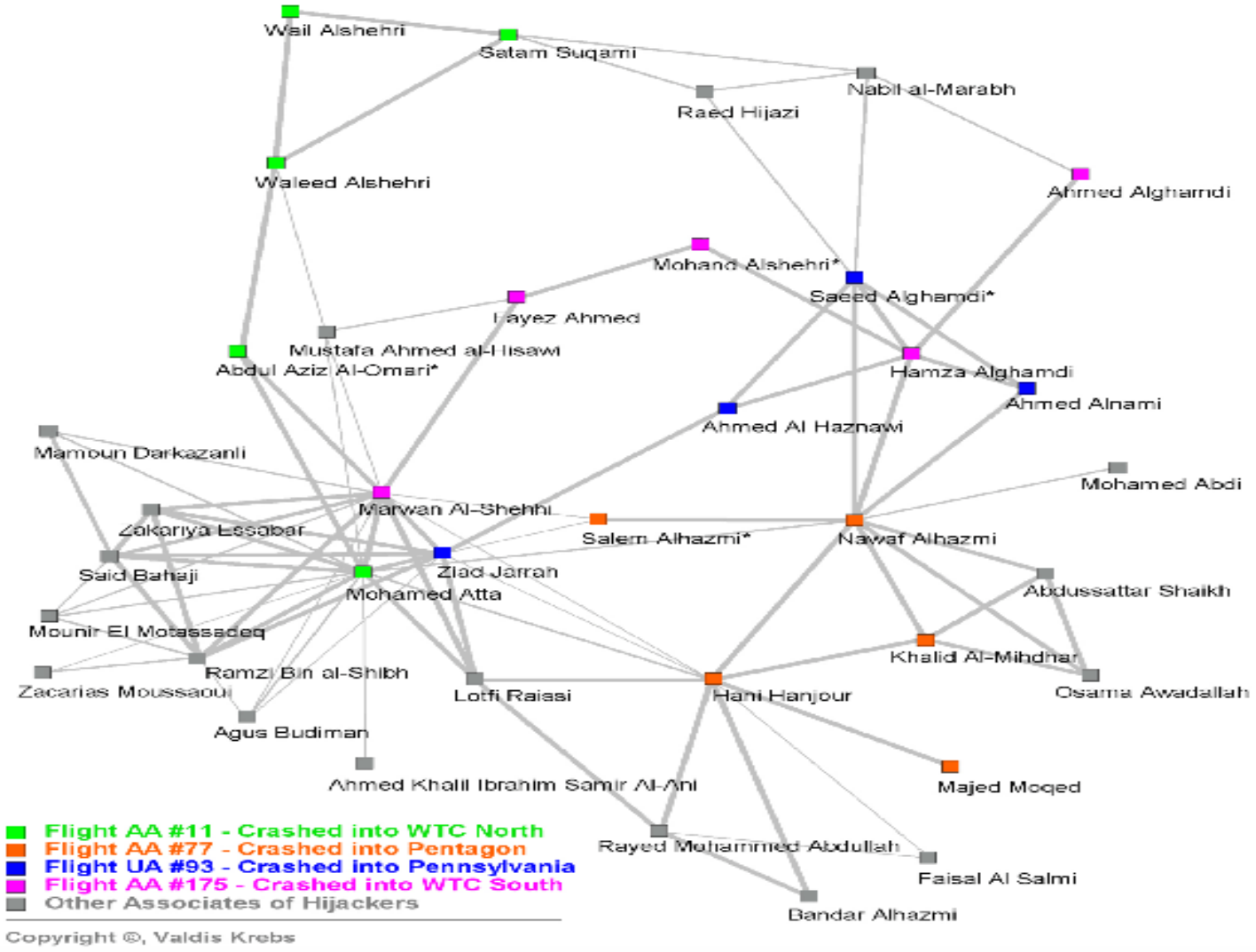


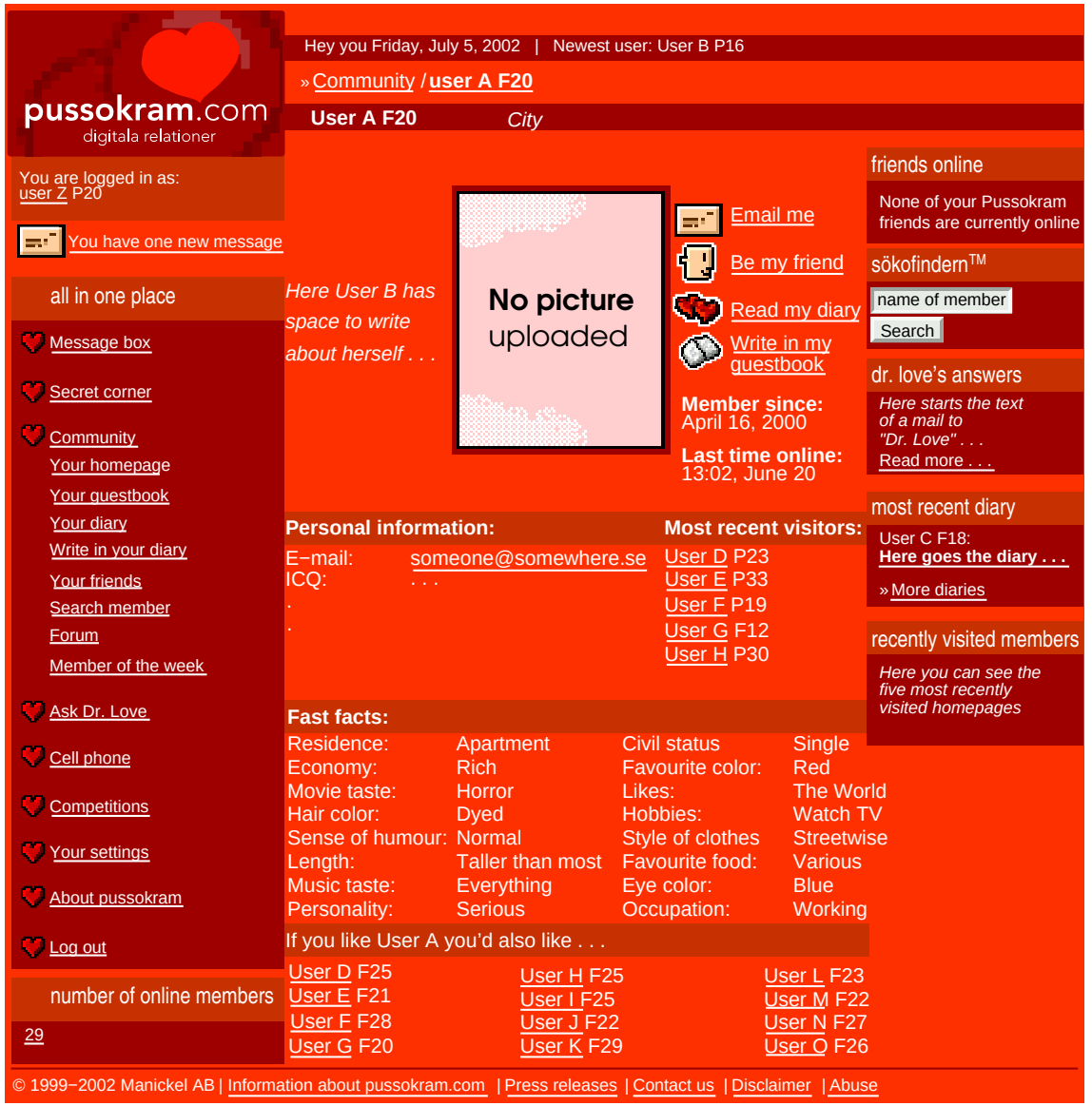
Figure 4. Hijacker's Network Neighborhood

WHY INTERNET COMMUNITIES

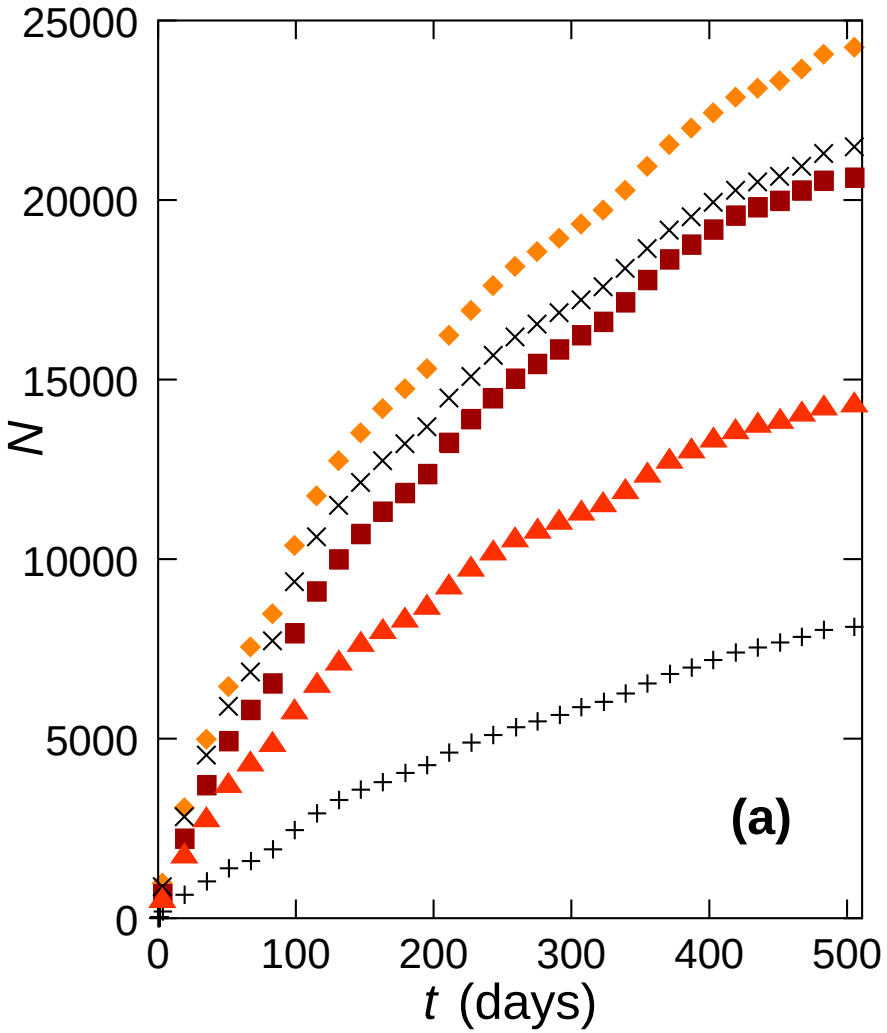
- large size (nice for statistical mechanics style approach)
- precise timing resolution
- formation more close to real acquaintance network than other methods
- interesting *per se* (I think)

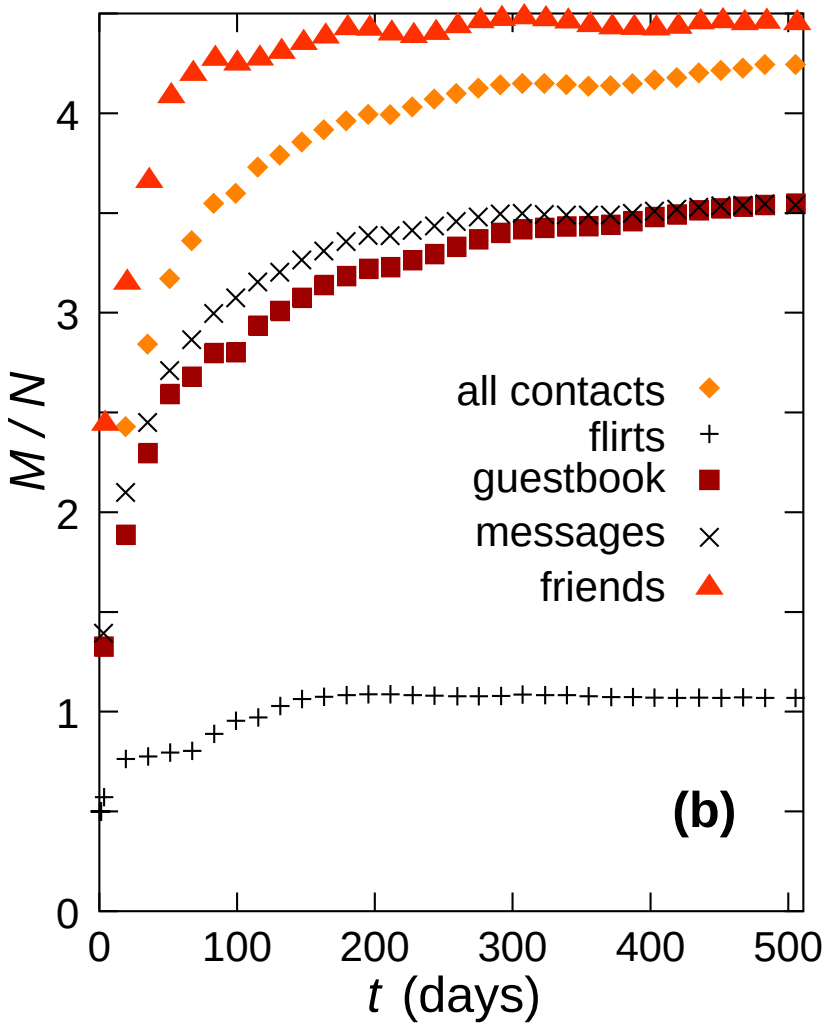
about pussokram.com:

- ~ 130 000 users
- ~ 70% female
- average age 21 yrs (?)



TIME DEVELOPMENT

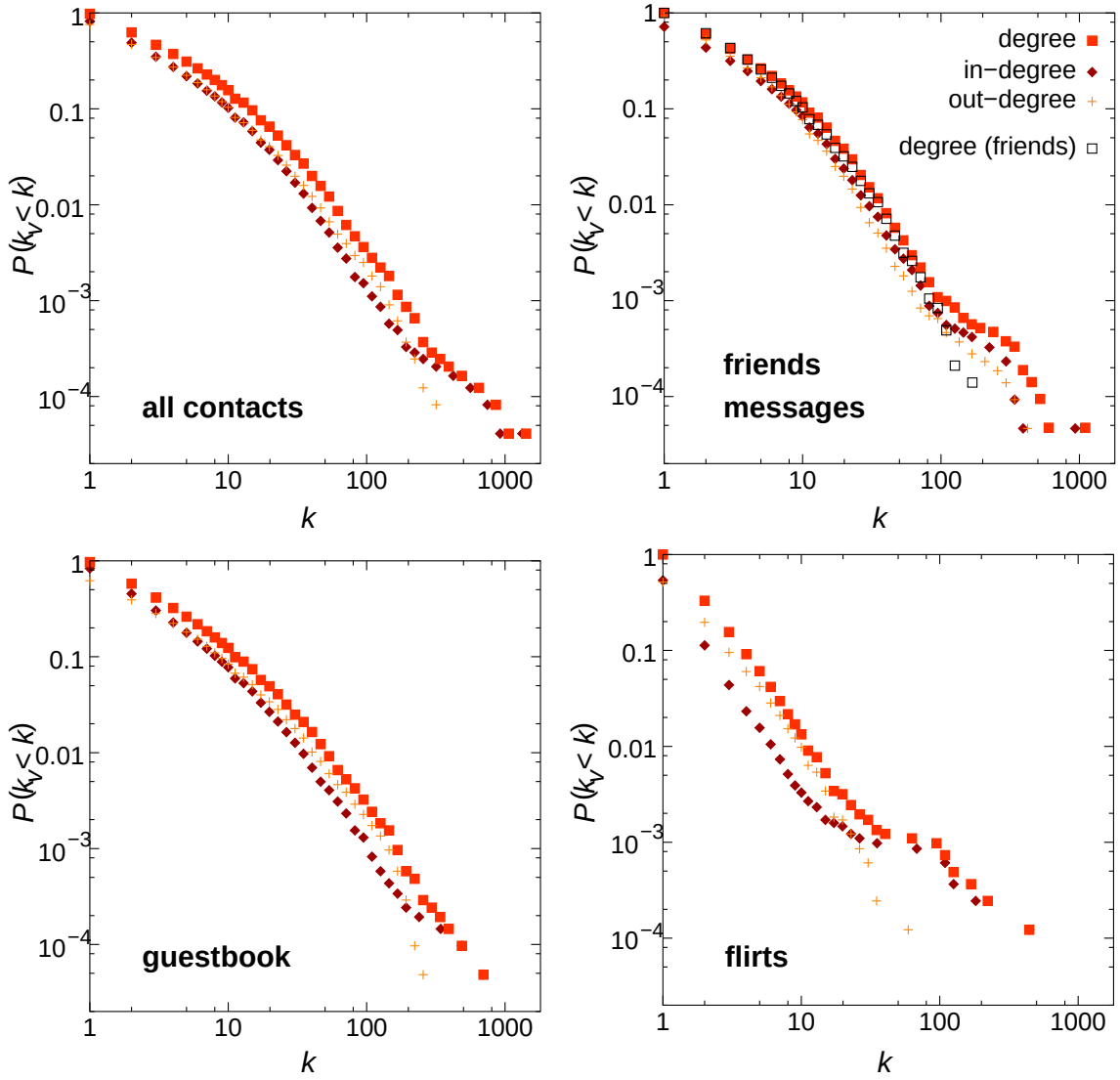




DEGREE DISTRIBUTION

. . . quite scale-free, as seen in e.g. web of sexual contacts.

(There has to be a cut-off but its nowhere to be seen)



Preferential Attachment
(Simon, Barabási, Albert)

A new node v attach to an old u with the probability

$$P_v(u) = \frac{k_u}{\sum_{w \in V} k_w}$$

Possible applications:

- World Wide Web
- Internet
- Collaboration networks in competitive fields

Pairwise introduction
(Davidsen, Ebel, Bornholdt)

Links between neighbors of vertices are formed at random. Nodes loose all their links from time to time.

- Regular acquaintance networks

Walking (Vazquez)

New nodes choose an old node and performs a random walk starting from this node. The random walk is terminated (at a node v) with probability p and an edge is added between the new node and v.

- Citation networks

Pairwise fitness (Caldarelli *et al.*)

1. All vertices are assigned a number x_v, from some

probability distribution. 2. An edge is drawn between a pair (v, u) of vertices with probability $f(x_v, x_u)$.

f is a symmetric function e.g. $\Theta(x_u + x_v - z)$ and the x_v does *not* have to be a power-law to get a power-law degree distribution.

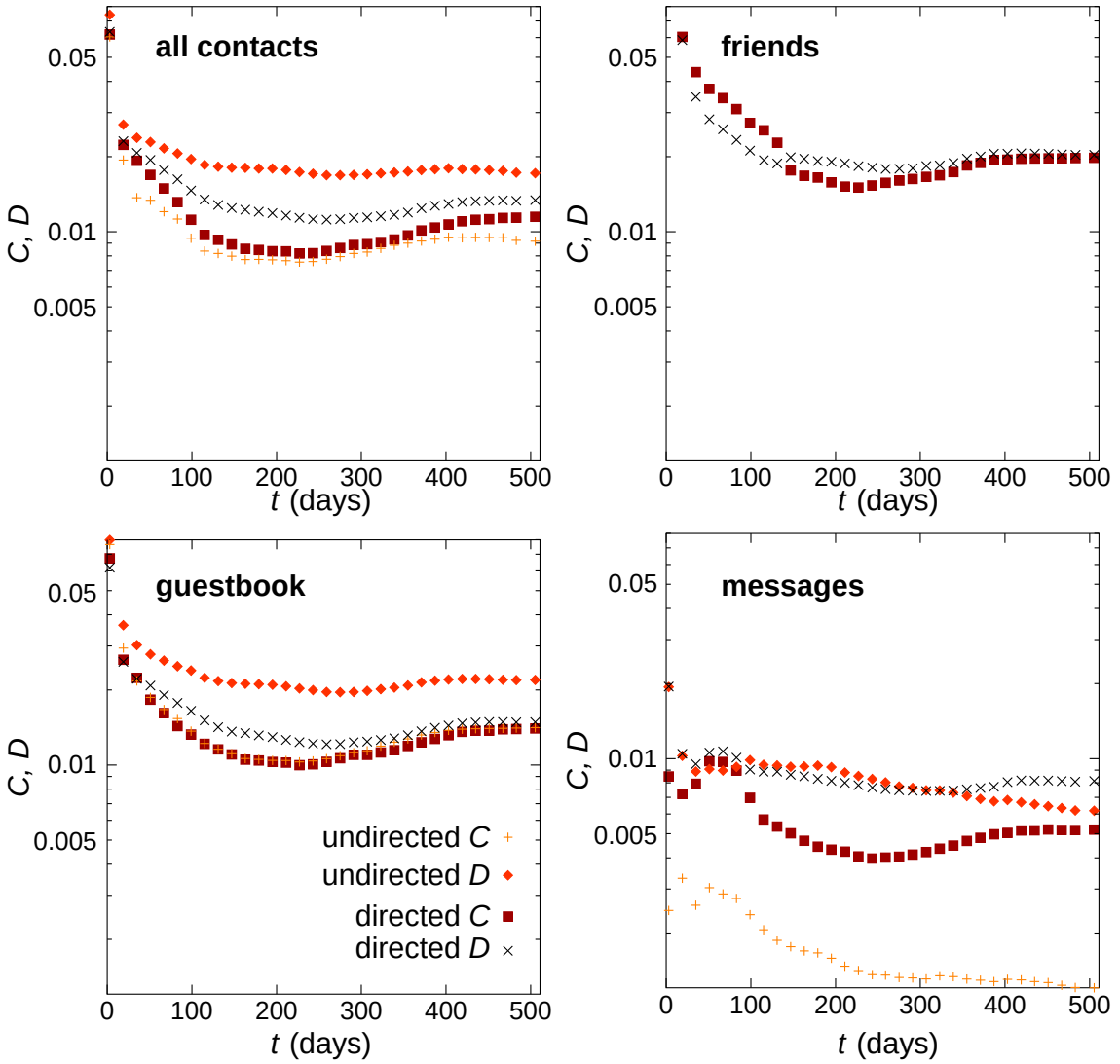
- WWW
- Internet communities (?)

DENSITY OF SHORT CIRCUITS

C—density of triangles,
D—density of squares.

High C in acquaintance networks.

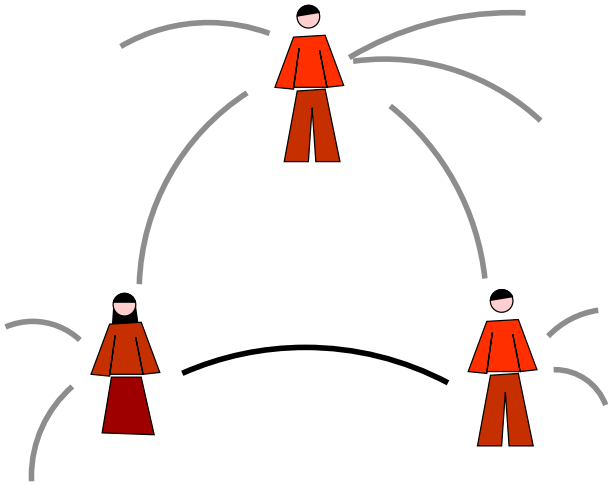
$C > D$ in other empirical network data. Here $D > C$ since most contacts are romantic & heterosexual.

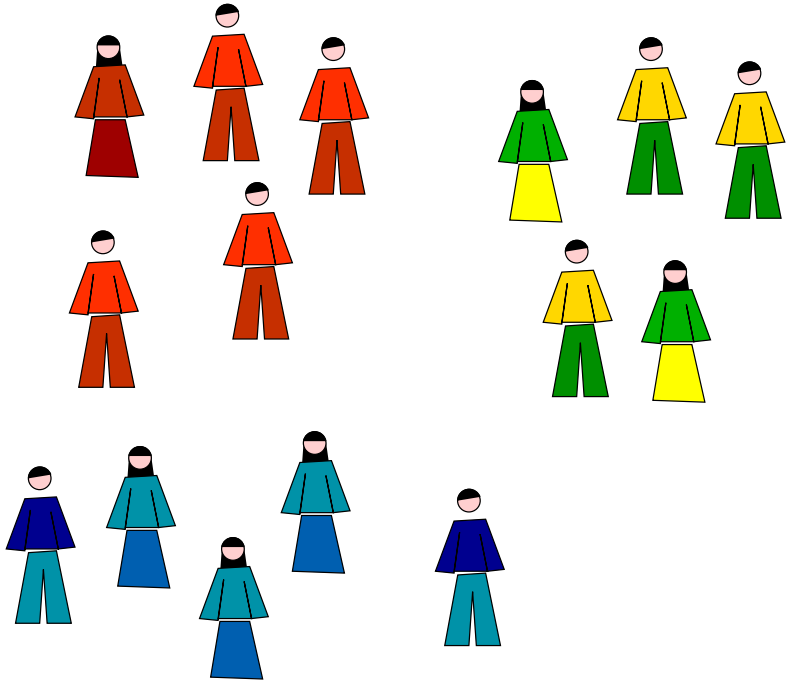


MECHANISMS FOR CLUSTERING

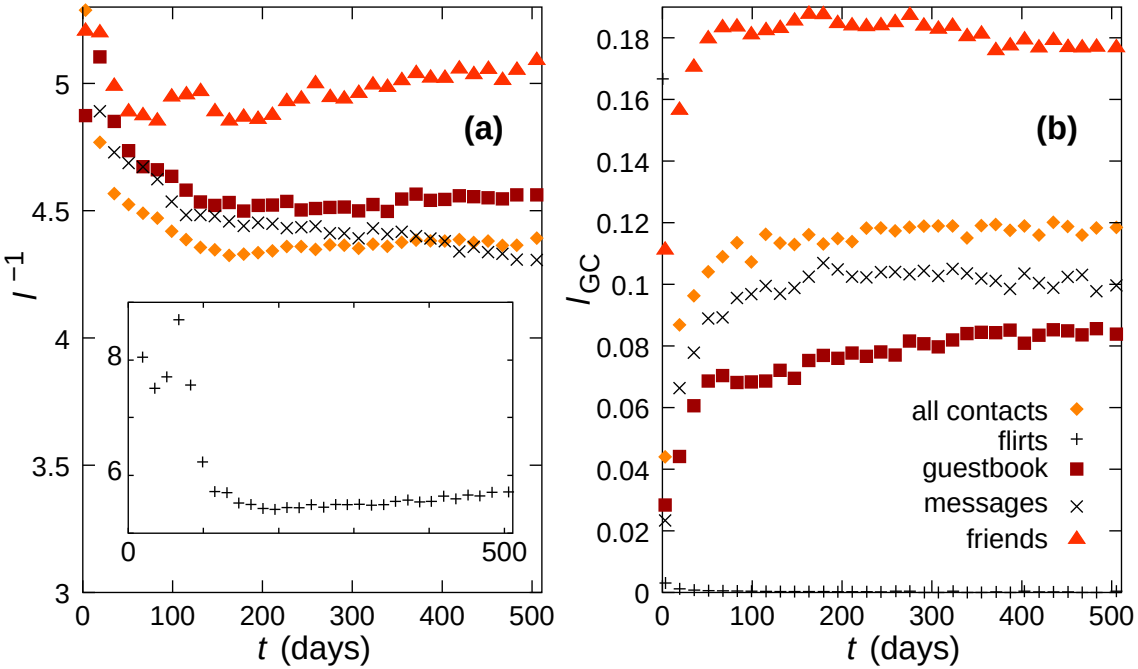
Introduction of friends. (Doesn't work for Internet communities)

Maybe hierarchical (community structure) gives large enough average degree in the clusters to get a finite clustering.



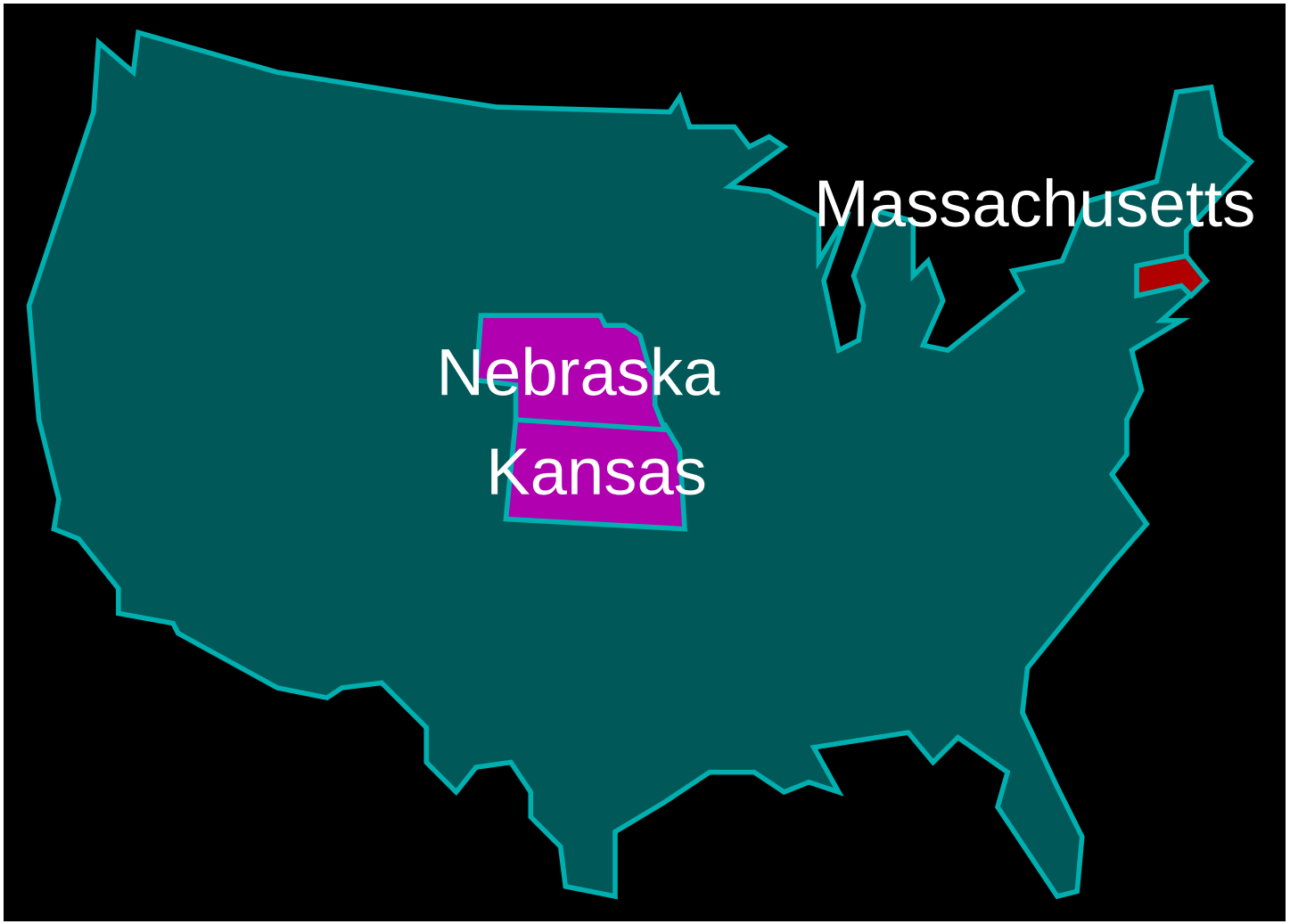


CHARACTERISTIC PATHLENGTH



$$l_{\mathrm{GSCC}} = \frac{1}{|A_{\mathrm{GSCC}}|} \sum_{(u,v)\in A_{\mathrm{GSCC}}} d(u,v)$$

$$l^{-1} = \frac{1}{M} \sum_{(u,v)\in A} \frac{1}{d(u,v)}$$



$\Rightarrow$ six handshakes between you and anyone else in the world . . .