New Complex Networks for Social Relations

Hans J. Herrmann

Computational Physics, IfB
ETH, Zürich, Switzerland

Marta C. González
José Soares Andrade Jr.
Pedro Lind
Herman Singer

International Workshop on Challenges and Visions in the Social Sciences
ETH Zürich
August 18-23, 2008
The statistical spreading of information or contacts in societies has many practical implications. It involves non-equilibrium phenomena where fluctuations and correlations play an important role.

Usually these processes are studied by sociologists using surveys including many parameters and obtaining qualitative results. Recently techniques from statistical physics and in particular complex networks and non-linear dynamics have been used to make simplified models and obtain quantitative laws.
Data from schools

Survey interviewing 90,118 students from 84 schools in US
(Add Health Program)

Visualization using "pajec" with T. Vicsek and J. Kertész
Schools

3-cliques

4-cliques
Schools

affinities between black and white students
Complex Networks

Food web of the North Atlantic Ocean
Scale-free networks

\[ P(k) \propto k^{-\gamma} \]

WWW:
\[
\begin{align*}
\gamma_{out} &= 2.4 \\
\gamma_{in} &= 2.1
\end{align*}
\]

Model: Barabasi-Albert \( \gamma = 3 \)

Internet
\( \gamma = 2.4 \)
actors
\( \gamma = 2.3 \)

HEP
\( \gamma = 2.3 \)
neuroscience
\( \gamma = 2.1 \)

scientific collaborations
Apollonian network

- scale-free
- (ultra) small world
- Euclidean
- space-filling
- matching
Apollonian packing
Other applications

• Force networks in polydisperse packings
• Highly fractured porous media
• Networks of roads
• Systems of electrical supply lines
Degree distribution

scale-free: \( P(k) \propto k^{-\gamma} \)

\[ W(k) \propto k^{1-\gamma} \]

\[ k \quad m(k, n) \]
\[ 3^n \quad 3 \]
\[ 3^{n-1} \quad 3 \times 2 \]
\[ 3^{n-2} \quad 3 \times 2^2 \]
\[ \vdots \quad \vdots \]
\[ 3^2 \quad 3 \times 2^{n-1} \]
\[ 3 \quad 3 \times 2^n \]
\[ 1 \quad 2^{n+1} \]

\[ \gamma = 1 + \frac{\ln 3}{\ln 2} \approx 2.585 \]

\( N_n = \) number of sites at generation \( n \)
\( m(k, n) = \) number of vertices of degree \( k \)

cumulative distribution \( W(k) = \int_{k>k}^{} m(k^*, n) / N_n \)
Small-world properties

Clustering coefficient:

\[ C = \frac{2}{k(k-1)} \times \text{number of connections between neighbors} \]

\[ C = 0.828 \]

Shortest path:

\[ l \propto (\ln N)^{3/4} \]
Percolation threshold

bond percolation
at $p_c : P_1, P_2$ and $P_3$ simultaneously connected

$p_c \propto L^{-\frac{1}{\nu}}, L = \sqrt[\uparrow]{N} \quad \nu \approx 3$

porous media \uparrow Archie's law

epidemics
random failure
Electrical conductance

- from center to all sites of generation $n$:
  \[ S_1 \{ S_2 \} = \text{number of outputs of each site} \]
  \[ n - 1 \rightarrow n \uparrow \quad S_1 \rightarrow 3 \times S_1, S_2 \rightarrow 2 \times S_2 \uparrow \quad z = \frac{2}{3} \]
  current distribution $P(I) \propto I^{-1.24}$

- from $P_1$ to $P_2$:
  $z \approx 0.45$
Critical temperature of Ising model

coupling constant $J_n \propto n^{-\alpha}$
correlation length $J_n$ diverges at $T_c$
free energy, entropy, specific heat are smooth
magnetization $m : e^{-\mathcal{H}_i} \quad T \to \Lambda$

opinion
Generalized Apollonians

\( m = 0 \)  
\( n = \infty \)

\( m = 1 \)

\( m = 2 \)

\( m = 3 \)

\( m = 4 \)

\( m = \infty \)
Fig. 3. Osculatory packing for white distribution of circle centers with an upper bound for allowed circle radii.
Apollonian packing in 3D

Tetrahedron
Conclusions on Apollonian networks

• Stable against random failure.
• Opinions stabilize at any interaction strength.
• Supply current distribution follows power-law while point-to-point distribution is log-normal.
• Apollonian networks are between small world \( l \propto \ln N \) and ultrasmall world \( l \propto \ln \ln N \).
• Force networks of polydisperse packings have few contacts between grains and their rigidity increases with density like a power-law.
• Porous medium follows Archie’s law. \( \Phi_\chi = 0 \).
Gossip
Gossip on network

[Diagram of a network with nodes and connections]
Gossip propagation

Spreading time $\tau$ is the minimum time it takes for a gossip to reach all accessible persons.

$n$ is the total number of persons that eventually get the gossip.

We also define the "spread factor" $f$:

$$f = \frac{n}{f}$$

where $k$ is the degree of the victim.
Gossip in schools
Gossip on random graph
\[ \tau = A + B \ln k \]

Barabasi-Albert

\begin{align*}
\text{m=5} & \\
\text{m=7} & \\
\text{m=3} & 
\end{align*}
Spreading time

schools

Apollonian

B-A, m=7
Spreading factor

(a) schools

(b) Barabasi-Albert network

$m=7$
Spreading factor

\[ k_0 \]

for Barabasi-Albert

\[ m=3 \]

\[ m=5 \]

\[ m=7 \]

Watts-Strogatz

\[ f \]

\[ C/C_0 \]

\[ \tau \]

\[ \log p \]
Distribution of $\tau$

$$P(\tau) = e^{\tau (1 - \gamma)/B}$$

Apollonian Schools

B-A

$m=7$
Gossip with probability $q$
Gossip about famous people

symbols = schools    lines = Barabasi-Albert
Degree distribution

\[ K_{nn}(k) \text{ is the average degree of the nearest neighbors of a site of degree } k. \]

(a) Average over all schools

(b) One school
Sexual contact networks

homosexual (Colorado Springs)  heterosexual (Sweden)

Agent model

soft-disks in 2d

interaction potential

\[ u(r_{ij}) = U_0 \left( \frac{\dot{c}}{c} \frac{\sigma}{r_{ij}} \right)^{12} \left( \frac{\dot{c}}{c} \frac{\sigma}{r_{ij}} \right)^6 \frac{\dot{\hat{r}}}{\hat{r}} + U_0 \], \quad r_{ij} \leq r_c = 2^{1/6} \sigma

collision time

\[ \tau_{\text{coll}} = \frac{1}{\rho} \frac{1}{2r_c} \sqrt{\frac{m}{\pi T k_B}} \]

Marta González
Evolution of network
Evolution of network with life-times
Comparison to schools
Shortest path and clustering coefficient

Comparison to schools

Degree distribution

Degree distribution: $K_{nn}(k)$ is the average degree of the nearest neighbors of a site of degree $k$.

Comparison to schools:
- one school
- average over all schools
3-clique communities

average over all schools

symbols: one school

full lines: agent model
SIS model of epidemic

\[ \Delta t_{inf} = (\text{MD time-steps}) \Delta t \]

\[ \lambda \propto \frac{\Delta t_{inf}}{\tau_{coll}} \]

\[ \lambda \propto \Delta t_{inf} \rho T^{1/2} \]

Directed percolation

Infection rate \( \lambda \)
SIS model for epidemics

\[ \lambda = 5.45 \]

Steady State

\[ \lambda = 1.2 \]
Data collapse

order parameter

\[ F_{IM}(\lambda) \approx \rho(t)_{\text{inf}} \lim_{t \to A_{\lambda}} \lambda > \lambda_c = \begin{cases} 0 & \text{for } \lambda < \lambda_c \\ 1 & \text{for } \lambda > \lambda_c \end{cases} \]

scaling law for \( \lambda > \lambda_c \):

\[ F_{IM}(\lambda) \approx 1 - \frac{1}{\lambda} \quad \text{with} \quad \lambda \equiv \frac{\Delta t_{\text{inf}}}{\tau_{\text{coll}}} \propto \Delta t_{\text{inf}} \rho T^{1/2} \]
Non-equilibrium phase transitions

System goes into the absorbing state
Influence of the density
Critical exponent $\beta$
Other critical properties

\( P(t) \) = probability that infection stays

\( n(t) \) = average number of infected agents
Critical exponents

\[ \delta_t = \frac{\log[P(t)/P(t/\Delta)]]}{\log\Delta} \]

\[ P(t) \propto t^{-\delta} \left(1 + \frac{a}{t} + \frac{b}{t^{\delta'}} + \ldots\right) \]

\[ \delta_t \propto \delta + \frac{a}{t} + \frac{\delta'b}{t^{\delta'}} + \ldots \]


### Critical exponents

<table>
<thead>
<tr>
<th></th>
<th>Mean Field</th>
<th>Moving Agents $\rho=0.1$</th>
<th>Moving Agents $\rho=0.2$</th>
<th>Contact Process 2D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_c$</td>
<td>1</td>
<td>1.0(8)</td>
<td>0.9(4)</td>
<td>1.6488(1)</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1</td>
<td>0.9(2)</td>
<td>0.7(7)</td>
<td>0.583(4)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>1</td>
<td>0.5(9)</td>
<td>0.5(3)</td>
<td>0.4505(10)</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0</td>
<td>0.1(5)</td>
<td>0.2(5)</td>
<td>0.2295(10)</td>
</tr>
<tr>
<td>$z$</td>
<td>1</td>
<td>1.3(0)</td>
<td>1.2(7)</td>
<td>1.1325(10)</td>
</tr>
</tbody>
</table>

$4\delta+2\eta=dz$ holds.
Networks properties

Rules for generating the network

- Each time one agent infects other a bond between the two is created.
- The infection lasts ($\Delta t_{inf}$ time steps), when one of the agents is recovered the bond disappears and the agent becomes susceptible to infection again.

Labeling of Cluster sizes

<table>
<thead>
<tr>
<th>Susceptible</th>
<th>RGB scale for cluster size</th>
</tr>
</thead>
</table>

- density = 0.1 $T = 1$, $\lambda = 2$
- density = 0.23 $T = 1$, $\lambda = 2$
Cluster size distribution

$N = 16$, $\rho = 0.1$, $\lambda = 10$

Second moment of the cluster size distribution

Same scaling as in percolation

Probability that an agent belongs to the biggest cluster

In Collab. with:
A.D. Araújo,
UFC, Brazil
Cluster numbers

Agents
\[ L = 32 \]
\[ \lambda = 6, 10, 23 \]

Percolation
\[ L = 32 \]
\[ p_c = 0.529 \]

\[ \tau \approx 2.05 \]
Clustering coefficient and degree distribution

\[ P(K) \]

\[ CC(\lambda) \]

\[ \rho=0.1, \lambda=3 \]
\[ \rho=0.4, \lambda=3 \]
\[ \rho=0.8, \lambda=3 \]
**Power-law distribution of infection time**

\[ P(\Delta t_{\text{inf}}) = (\gamma - 1)\Delta t_{\text{inf}}^{-\gamma} \quad 2 < \gamma \leq 3 \quad \Delta t_{\text{inf}} > 1 \]

\[
\lambda_c = \frac{\langle t_{\text{inf}} \rangle}{\tau_{\text{coll}}} = \frac{\gamma - 1}{\gamma - 2} \rho_0 2r_0 \sqrt{\frac{\pi Tk_B}{m}} > 1
\]

\[ \rho_c \bigg|_{\text{calc}} = 0.00718 \quad \rho_c \bigg|_{\text{num}} = 0.006(5) \]

\[ \beta = 2.6(5) \]
Conclusions on epidemics

✓ Novel effects are observed studying the SIS model of infection on a system of moving agents. A continuous range of critical exponents is found as a function of the density number of agents.

✓ Geometrical properties like the degree distribution and the clustering coefficient of the network of infected agents depend on the rate of infection and not on the density of agents. The distribution of cluster sizes is not a power law at the transition to spreading.

✓ Introducing a power law distribution of infection times the epidemic threshold becomes zero, but there is still a critical rate of infection that depends on the exponents of the distribution and the mean interaction time among the agents.
Sexual contact networks

homosexual (Colorado Springs)  heterosexual (Sweden)

Comparison with real data

J.J. Potterat et al,

Velocity of agent proportional

Barabasi-Albert scale-free network
Bipartite growing network
Comparison with real data

**homogeneous:**

- Real data
- Our model
- Previous model

**heterogeneous:**

- Real data
- Our model
- Previous model
Sexual contact networks

Degree distribution

Sexual contacts with any agent.

Sexual contacts only with uninitiated agents.
Sexual contact networks

**homosexual**  (Colorado Springs)  **heterosexual**  (Sweden)

Sexual contact networks

heterosexual

58% females
42% males

Sexual contact networks

\[ P(K) \]

\[ N_{\text{ntw}} = 2 \times 10^4 \]

\[ N_{\text{ntw}} = 10^4 \]

\[ N_{\text{ntw}} = 10^3 \]

\[ \gamma_t = 1 \]

\[ \gamma_h = 3 \]
Another friendship model

• Goal
  – friendship network model (spatially independent) for a fixed setting (e.g. enrolling at a university)
  – reproduce experimentally measurable quantities
  – natural emergence of community structures

Herman Singer, ETH
## Properties of the agent

- **agent properties**
  - **affinity** \( a_i \in [0,1] \)
  - list of past contacts, length \( n_i \)
  - maximal acquaintance parameter \( \lambda_i \)
  - absolute importance
  - relative importance

\[
\prod_i = \frac{k_i}{N \sum_{j=1}^N k_j}
\]

- behavior
  - agent can choose between meeting new contacts and meeting again already established contacts
  - similar to aging but important difference:
    - agent can accept new contacts throughout the simulations
    - at the expense of dropping old ones

\[ p_{ij} = \frac{f_{ij}}{n_i \sum_{k=1} f_{ik}} \]

→ every agent optimizes its interest
→ local, self organized community structure emerges
Friendship in formulas

- friendship is written as the weighted sum of the contributions
  - number of times $n_{ij}$
  - interest/affinity $a_i, a_j$

$$f_{ij} = \gamma f_1(|a_j - a_i|) + (1 - \gamma) f_2(n_{ij})$$

- with $\gamma > 0.8$

friendship functions:
- affinity $f_1$

$$f_1(|a_j - a_i|) = \left(1 - |a_j - a_i|\right)^\kappa$$

- with $\kappa \geq 2$, here 2,3,5,10

- frequency $f_2$
  - exponential

$$f_2(n_{ij}) = 1 - e^{-\lambda f n_{ij}}$$

- with $\lambda_f = 0.2 \leftrightarrow$ saturation on $\sim 25$ contacts
Algorithm

- $t=0$: initialize $N$ agents, no connections with parameters $a_i$, $n_i=0$, $\lambda_i$
- $t \rightarrow t+1$:
  - choose an agent randomly: $i$
  - decide on contact mode with probability $p = 1 - e^{-\lambda_i n_i}$
    - select an agent from pool with probability $\Pi_j$
      - add to contact list
    - otherwise select an agent in contact list with probability $p_{ij}$
- prevent social isolation and take into account random encounters:
  - introduce threshold $p \geq \theta_p$
Degree distribution

\[ P(k) \]

- Gonzalez et al.
- Exponential
- Poisson
- experimental
- 10000 x N = 1000
- 500 x N = 10000
• \( P(k|k') \): probability at a node with \( k \) to find a neighbor with \( k' \)

\[
K_{mn}(k) = \sum_{k'=1}^{N} P(k | k') k'
\]
Clustering coefficient
Friendship distribution

- Experimental results
  - 417 high school students
  - "who are your best friends?"
  - probability that a person is mentioned n times
Conclusion

- Social relation is difficult to quantify.
- Many simplified models are possible.
- Parameters can be tuned to agree with some data.
- One can find general conclusions.
- But does anybody believe these generalities?
- Does anybody care about this research?
- Do we need socio-physics?
- What does one learn from modeling?