SOCIAL SYSTEMS (as not only) COMPLEX INTERACTIVE NETWORKS

Péter Érdi\textsuperscript{1,2} and Gábor Csárdi\textsuperscript{3,1,2}

\textsuperscript{1}Center for Complex Systems Studies, Kalamazoo College, Kalamazoo, Michigan


\textsuperscript{3}Computational Biology Group, Department of Medical Genetics, Université de Lausanne. Lausanne, Switzerland
1. The place of network models is social dynamics

2. The direct and inverse problem of evolving networks

3. Stochastic kinetic modeling of the US patent citation network
Center for Complex Systems Studies

Budapest CNS Group
1 The place of network models is social dynamics

**bulk models**: spatially homogeneous temporal process.

Growth models with a single variable:

\[
\frac{dx}{dt} = x^m, \quad x(t) = x(0) \left( \frac{t_c - t}{t_c} \right)^{-\frac{1}{m-1}},
\]

explosion (finite-time singularity) with \( m > 1 \),
while \( m = 0 \) leads to linear, and \( m = 1 \) to exponential growth

Increase at CONSTANT (i.e. state-INdependent) RATE: no feedback
Increase PROPORTIONAL to Actual State (linear positive feedback)
Increase with with HIGHER-THAN-LINEAR positive feedback
Temporal dynamics: interacting variables: equilibrium, oscillation, chaos

A) war dynamics

principle of competitive exclusion vs. “permanent war”,
Gause, Strogatz, Epstein
B) drug propagation and control

(adapted from Behrens et. al, 2004)
C) economic cycles and chaos

Business Cycles: Kaldor model

\[ k = k_e \]

\[ \frac{dy}{dt} = 0 \]

\[ \frac{dk}{dt} = 0 \]

\( y \): income, \( k \): capital, \( I \): investment, \( S \): saving.

Modified versions with time delay.
Temporal dynamics: interacting variables: equilibrium, oscillation, chaos

D) biological and social epidemics

\[ S(t) \text{ – Susceptible} \]
\[ I(t) \text{ – Infectious} \]
\[ R(t) \text{ – Recovered} \]
Temporal dynamics: interacting variables: equilibrium, oscillation, chaos

A) war dynamics

B) drug propagation and control

C) economic cycles and chaos

D) biological and social epidemics
Spatiotemporal processes

closed systems: homogenization
open systems: pattern formation

spatial: as urban (racial) segregation

- propagation of matter, energy: Physics, virus: Biology
- propagation of matter: goods, money etc.: Economics
- propagation of drugs: Criminology
- propagation of “information” (knowledge, opinion, gossip, fear, trust) among humans: Behavioral Social sciences

Real world models: **coupling!!**
Joshua Epstein: Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations
Social Networks

- microscopic – discrete objects connected by relationship (static) or flows (dynamic)

- nodes: people, organizations, communities, documents

- propagation on networks (threshold vs. non-threshold phenomena)

- evolution of networks

of course: not only social networks
example for on: propagation of “information” among computers: malware (malicious software)
example for of: ontogeny, plasticity and regeneration of the neural networks

More about evolving networks !!
School Friendship Network

http://www-personal.umich.edu/~mejn/networks/
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Complex Systems as Evolving Networks
Network evolution means that the structure changes in time:

- nodes/edges are added and/or
- nodes/edges are deleted.
Network evolution depends on

- structural properties of the nodes/edges, e.g. node degree,

- intrinsic (non-structural) properties of the nodes/edges, e.g. node color,

- other (not node-specific) parameters.
The direct and inverse problem of evolving networks

Kernel Function

- **Property vector**: \((\mathbb{X}), x\)-vertex: a vertex with property vector \(x\)
- **Kernel function**: \(A : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}\)
- **Higher kernel function value \(\rightarrow\) more probable the realization of the given edge**
- **The probability that a given edge \(e\) connects an \(x\)-vertex to a \(y\)-vertex**: 

\[
P[c(e, x, y) = 1] = \frac{A(x, y)N(t(e), x, y)}{\sum_{(x', y') \in \mathbb{X} \times \mathbb{X}} A(x', y')N(t(e), x', y')}
\]

\(c(e, x, y)\) are indicator random variables, (one for every edge–property vector triple. 
\(c(e, x, y) = 1\) if and only if edge \(e\) connects an \(x\) vertex to \(y\)-vertex; \(t(e)\) is the time step before the addition of edge \(e\); \(N(t(e), x, y)\) is the number of possible \(x\)-\(y\) connections in time step \(t(e)\). \]
Citation networks

- special directed networks:
  - edges and vertices are never deleted from the network
  - all outgoing edges of a vertex are added right after the vertex itself
  - we will assume that a single vertex is added to the citation network in each time step.
  - they contain no loops
Figure 3.1: Some snapshots for a citation network. Since the new vertices are always added to the right and to the top of old vertices, all edges go to the left or downwards. Here we show three snapshots, vertices 6-10 are added between the first and the second and 11-19 between the second and the third. Notice that all outgoing edges of a vertex are added with the vertex itself.
Direct and Inverse Problems

Kernel function(s) $A(\cdot, \cdot)$ → Kernel-based generator → Network (artificial)

Kernel function(s) $A(\cdot, \cdot)$ → Kernel-based measurement → Network (real)
Direct problems

- The difficulties in social science: from Kepler to Newton??
- Kepler’s laws: data-driven, inductive, global, integral.
- Newton’s laws: data-motivated and concepts driven, deductive, local, differential.
- The *a priori* knowledge of the kernel functions, which basically governs the dynamic evolution of the systems, is difficult: social dynamics is in the post-Keplerian and pre-Newtonian age (and maybe it remains there).
Solving the Inverse problem

The Frequentist Method

\[
A(x) = \frac{P[c(e, x) = 1]S(t(e))}{N(t(e), x)}
\]

How to measure? (approximate?)

An observed edge \( e \) either cites an \( x \)-vertex or not:
if not: the correct estimation for \( P[c(e, x) = 1] \) is zero, thus the approximation, the score for \( A(x) \) is also zero.
if yes: the approximation is

\[
\bar{A}_e(x) = \frac{S(t(e))}{N(t(e), x)},
\]

where \( \bar{A}_e(x) \) reads as “the approximation of \( A(x) \) based on edge \( e \)”. Taking the average of the scores we get

\[
\bar{A}(x) = \frac{1}{|E_x|} \sum_{e \in i(x)} \frac{S(t(e))}{N(t(e), x)},
\]

where \( |E_x| \) is the number of time steps in which there was at least one \( x \)-vertex present in the network.

There is an iterative approach to get a consistent solution for \( A(\cdot) \), and the convergence is proven (G. Csárdi).
Solving the Inverse problem
The Maximum Likelihood Method

The goal is to extract a kernel function from the network evolution data.

The function to be maximized is the probability that a kernel function generates exactly the observed network. I.e. (for citation networks):

\[
\prod_e \frac{A(x_e)}{S(t(e))} = \prod_{i=1}^n A(i)^M_i \prod_e \left[ \sum_{i=1}^n p_i^{t(e)} A(i) \right]^{-1}.
\]

The \( S \) normalization factor is

\[
S(t(e)) = N^{t(e)} \sum_{i=1}^n p_i^{t(e)} A(i).
\]

(\( N^{t(e)} \) factors can be neglected.)

again: existence and uniqueness were proved by GCs, and the whole procedure has been generalized for non-citation networks.
3 Stochastic kinetic modeling of the US patent citation network

collaboration with Katherine Strandburg and Jan Tobochnik

WHY STUDY IT?

- Very large network (about 4 million patents, between 1975 and 2005)
- Data available electronically (NBER dataset + USPTO)
- Comparison to scientific citation networks
- Relevance to patent policy
- Theory of growing networks can be applied

How to define a mathematical model framework for describing the temporal evolution of the patent network?

How many state variables we need?
How many state variables we need? *In-Degree (only)*

An in-degree dependent kernel function can be very well fitted with:

\[ A(d) = d^\alpha + a; \alpha \text{ exponent is close to unity; (may lead to scale-free networks)} \]

Only tells that this is the best form to come up with *if* we want to model the network based on in-degree only. In other words, if we add additional vertex properties to the model, we might get a better model.
How many state variables we need? *In-Degree and Age*

\[ A(d, l) = A_d(d) A_l(l) : \text{linear preferential attachment times double Pareto age-dependent part} \]

Fig. 6: Sections from the *in-degree* and *age* based maximum likelihood fitted kernel function for the US patent citation network. Both plots have logarithmic axes.

\[ A_l(l) = \begin{cases} 
(l/t_p)^{\beta_p-1} & \text{if } l \leq t_p, \\
(l/t_p)^{\alpha_p-1} & \text{if } l > t_p. 
\end{cases} \]
How many state variables we need? In-Degree, Age, Cited Category, Citing Cat.

more than 400 patent classes and six big categories:

Chemical, Computers and Communications, Drugs and Medical, Electrical and Electronic, Mechanical and Others: patent citation network is highly assortative (Chemical cites Chemical)

• \( A^c(d, l, p) = c_p \times A^c_d(d)A^c_l(l) \) was fitted for each \( c \) citing vertex type

• \( A^c_d(d) \) has a preferential attachment form

• \( A^c_l(l) \) is the double Pareto function

• \( c_p \) are category-dependent constants

• separate kernel functions for different citing categories \( \rightarrow \) different parameters
Interpreting the kernel function (*In-Degree*, Age, *Cited Category*, *Citing Cat.*)

• The categories are really different.

• Stratification is higher in some categories than in others, i.e. *Computers & Communication* and *Drugs and Medical*.

• In these categories “a few” highly important patents receive the majority of the citations.

\[
A^c(d, l, p) = c_p \times A^c_d(d) A^c_l(l)
\]

\[
A^c_d(d) = d^{\alpha_c} + a_c
\]
Interpreting the kernel function \((\text{In-Degree}, \text{Age}, \text{Cited Category, Citing Cat.})\)

\[
\begin{align*}
A^c(d, l, p) &= c_p \times A_d^c(d) A_l^c(l) \\
A_l^c(l) &= \begin{cases} 
(l/t_p)^{\beta_p^c - 1} & \text{if } l \leq t_p^c, \\
(l/t_p)^{-\alpha_p^c - 1} & \text{if } l > t_p^c.
\end{cases}
\end{align*}
\]

- Patents in different categories to out-of-date at a different rate.

How may state variables do we need? Which one is the right model?

- The goodness function:

\[
\frac{1}{|E|} \left[ \sum_e \log \frac{A(x_e)}{S(t(e))} - \sum_e \log \frac{1}{t(e)} \right].
\]

- The probability that the kernel function generates the observed network.

- Almost,

  1. we take the logarithm of it and
  2. subtract the probability that the completely random kernel function generated the network.
  3. Plus, we divide by the number of edges to make it size-independent.
How may state variables do we need? Which one is the right model?

<table>
<thead>
<tr>
<th>Properties</th>
<th>Form</th>
<th>Goodness</th>
</tr>
</thead>
<tbody>
<tr>
<td>in-degree</td>
<td>free</td>
<td>0.2431</td>
</tr>
<tr>
<td>in-degree</td>
<td>$d^\alpha + a$</td>
<td>0.2394</td>
</tr>
<tr>
<td>i-d, age</td>
<td>free</td>
<td>0.4541</td>
</tr>
<tr>
<td>i-d, age</td>
<td>double Pareto</td>
<td>0.4373</td>
</tr>
<tr>
<td>i-d, citing category</td>
<td>free</td>
<td>0.2583</td>
</tr>
<tr>
<td>i-d, citing category</td>
<td>$d^\alpha + a$</td>
<td>0.2555</td>
</tr>
<tr>
<td>i-d, cited category</td>
<td>free</td>
<td>0.2656</td>
</tr>
<tr>
<td>i-d, cited category</td>
<td>$c_{\text{cat}}(d^\alpha + a)$</td>
<td>0.2597</td>
</tr>
<tr>
<td>i-d, cited, citing</td>
<td>free</td>
<td>1.1744</td>
</tr>
<tr>
<td>i-d, cited, citing</td>
<td>$c_{\text{cat}}(d^\alpha + a)$</td>
<td>1.1664</td>
</tr>
<tr>
<td>i-d, age, cited, citing</td>
<td>$c_{\text{cat}} \times \text{double Pareto}$</td>
<td>1.3454</td>
</tr>
<tr>
<td>forest fire model</td>
<td>free</td>
<td>2.1034</td>
</tr>
<tr>
<td>i-d, forest fire</td>
<td>free</td>
<td>2.2539</td>
</tr>
</tbody>
</table>
Time-dependent rules: change in the dynamics

- Single kernel during the whole network development? not necessarily

- Legal changes in 1991

- Lerner and Jaffe describe the recent history of the patent system, in which patents were both more easily obtained and potently enforced. They believe that the patent review process of the U.S. Patent and Trademark office should be more rigorous and that the playing field between litigants should be more level. The protection for true innovators created by a workable patent system is vital to technological change and economic growth, they write.

- What does our model tell us about 1991?
Time-dependent rules: change in the dynamics

We use a sliding time window of 500,000 patents to measure the time dependent kernel-function.

Fig. 8: Change of the $\alpha$ exponent in the US patent network using two models. The left plot is the maximum likelihood fit of the in-degree dependent model $A(d) = d^\alpha + a$. The right plot is based on an in-degree and age dependent model.

Increase in the exponent of preferential attachment.
Compare patents and scientific citations

- What if the phenomenon is not specific to the patent network? E.g. electronic patent databases just appeared around 1990.

- Then the same should have happened in scientific citation networks as well.

- The APS network. Scientific citations among the journals of the American Physics Society, 110 years, up to 2005.

- About 400,000 papers and 3.5 million citations.
• Interestingly, the same forms are good.

• E.g. in-degree model, preferential attachment form:
Compare patents and scientific citations

- in-degree and age dependent model, in-degree part, prefential attachment:
Compare patents and scientific citations

- **in-degree** and **age** dependent model, age part, double-Pareto form:
• Preferential attachment exponent, in-degree and age dependent model:
• Aging exponent, in-degree and age dependent model:
Compare patents and scientific citations, time dependence

- Variation is rather small, compared to the patent network.

- Preferential attachment exponent:

- No sign of the same increased stratification as in the patent network.
Lessons learned from network analysis

- network structure strongly influences the diffusion of information on it

- “stratification” – more and more nodes with very few citations and less and less nodes with many citations

- “sleeper patents” matter: it may happen that old patents gain new significance in light of later advances

- changes in the laws of the patent review process and in the level of rigorousness of the patent examinations over-accelerated the process.

- technological progress is somewhat unpredictable
Prediction in social science/politics is in any case: 

difficult
Prediction in social science/politics is in any case: difficult
THANK you for your attention