Explosive percolation in random graphs

Raissa D’Souza
University of California, Davis
Dept of Mech. and Aero. Eng., Dept of CS
Complexity Sciences Center
External Professor, Santa Fe Institute
Networks are ubiquitous:

Networks:
- Transportation Networks/Power grid (distribution/collection networks)
- Biological networks
  - protein interaction
  - genetic regulation
  - drug design
- Computer networks
- Social networks
  - Immunology
  - Information
  - Commerce

A collection of interacting networks
Modeling networks as random graphs


- Configuration models (Bollobás 1980, Molloy and Reed RSA 1995). Enumerating over all networks with specified \( \{p_i\} \).


- Growth by copying (Kumar, Raghavan, Rajagopalan, Sivakumar, Tomkins, Upfal FOCS 2000), including duplication/mutation (Vazquez, Flammini, Maritan, Vespignani, ComPlexUs 2003)

- Many more . . .
Building a random instance of a network


- Start with $N$ isolated vertices.
- Add random edges one-at-a-time.
  \[ E = \frac{N(N-1)}{2} \]  
  total edges possible.
- After $\mathcal{E}$ edges, probability $p$ of any edge is
  \[ p = \frac{\mathcal{E}}{E} = \frac{2\mathcal{E}}{N(N-1)} \]

What does the resulting graph look like?  
(Typical member of the ensemble)
\[ N = 300 \]

\[ p = \frac{1}{400} = 0.0025 \]

\[ p = \frac{1}{200} = 0.005 \]
Emergence of a “giant component”

- $p_c = 1/N$.
- $p < p_c$, $C_{\text{max}} \sim \log(N)$
- $p > p_c$, $C_{\text{max}} \sim A \cdot N$

(Ave node degree $t = pN$ so $t_c = 1$.)

Branching process (Galton-Watson); “tree”-like at $t_c = 1$. 
Phase transitions

Abrupt change in fundamental property of a system in response to slight change in controlling variable

- \( T < T_c \), **discontinuous jump** in thermodynamic quantity (e.g., Volume, \( V \))
- \( T = T_c \), **continuous change** in \( V \), but derivatives (e.g., \( \frac{\partial V}{\partial P} \)) diverge.
- (For \( T > T_c \), “supercritical fluid” state)
Universality classes I

Discontinuous – “first order”

- Phase coexistence / Latent heat
- Finite length scales
Universality classes II: Continuous – “second order”
Scaling behaviors

- Diverging correlation lengths and response functions.

- Heat capacity: \[ C_v = \left. \frac{\partial E}{\partial T} \right|_V \sim |T - T_c|^{-\alpha} \]

- Isothermal compressibility: \[ \kappa_T = -\frac{1}{V} \left. \frac{\partial V}{\partial P} \right|_T \sim |T - T_c|^{-\gamma} \]

- Magnetic susceptibility: \[ \chi = \left. \frac{\partial m}{\partial h} \right|_T \sim |T - T_c|^{-\gamma} \]

“Mean field” Ising and van der Waals gas \( \rightarrow \gamma = 1 \).

(Thermodynamic properties depend only on a small number of features –
dimensionality, symmetry – insensitive to underlying microscopic properties)
Erdős-Rényi – second order phase transition

• \( t < 1, \; C_{\text{max}} \sim O(\ln n) \)
• \( t = 1, \; C_{\text{max}} = n^{2/3} \)
• \( t > 1, \; C_{\text{max}} \sim An, \; \text{with} \; A > 1 \)

• The critical window

  \[
t = 1 + \lambda n^{-1/3} \quad \text{(where} \; t = 2e/n \text{)}
\]

• Mean field critical exponents

  \[
  \chi \sim (t_c - t)^{-\gamma}, \; \text{with} \; \gamma = 1.
  \]

  where \( \chi \) is the expected size of the component to which an arbitrarily chosen vertex belongs.
Erdős Rényi random graph: A continuous phase transition

\[ C_{\text{max}}(t_c - \epsilon) = N^{2/3} \]

\[ C_{\text{max}}(t_c + \epsilon) = N^{2/3} + A \]

\[ A = N^{-1/3} \]

\[ \lim_{N \to \infty} N^{-1/3}, A \to 0 \]

\[ C_{\text{max}}(t_c + \epsilon) - C_{\text{max}}(t_c) \to 0 \]
Connectivity – good or bad?

- Communications, Transportation, Synchronization, ...

versus

- Spread of human or computer viruses
Can any **limited perturbation** change the phase transition?

[Bohman, Frieze, *RSA* 19, 2001]

[Achlioptas, D’Souza, Spencer, 2009]

- Possible to **Enhance** or **Delay** the onset?

- The **“Product Rule”**
  - Choose *two* edges at random each step.
  - Add only the desirable edge and discard the other.

- **The Power of Two Choices**
  Azar; Broder; Mitzenmacher; Upfal; Karlin;
ProdRule: Explicit example

- Prod $e_1 = (7) \times (2) = 14$
- Prod $e_2 = (4) \times (4) = 16$

- To *enhance* choose $e_2$. To *delay* choose $e_1$. 
Product Rule

- **Enhance** – similar to ER but with earlier onset.

- **Delay** – changes from *continuous* to *discontinuous* transition!
Delayed Product Rule: Discontinuous change

\[ C_{\text{max}}(t_c - \epsilon) = N^{\gamma}, \text{ with } \gamma < 1 \]

\[ C_{\text{max}}(t_c + \epsilon) = A \cdot N \]

A increases with \( N \)!

\[
C_{\text{max}}(t_c + \epsilon) - C_{\text{max}}(t_c) \to A
\]

\[ ε = 0.0005 \]

\[ A = 0.001 N^{1/3} \]
The window $\Delta$ from $n^{1/2}$ to $0.5n$

- Let $e_0$ denote the last edge added for which $C_{max} < n^{1/2}$.
  (Recall ER has $n^{2/3}$ at $p_c$.)
- Let $e_1$ denote the first edge added for which $C_{max} > 0.5n$.
- Let $\Delta = e_1 - e_0$.

PR $\Delta \sim n^{2/3}$

ER (and BF) $\Delta \sim n$.

PR From $n^{1/2}$ to $0.5n$ in number of edges that is sublinear in $n$. 
Bounding $t_c$, where $t = e/n$

(Note, for ER, $t_c = 1/2$)

- For $t < t_c$, $C_{\text{max}} < n^{1/2}$.
- For $t > t_c$, $C_{\text{max}} > 0.5n$.

Jumps “instantaneously” from $C_{\text{max}} = n^{1/2}$ to $0.5n$. 
“Explosive Percolation in Random Networks”
From \( n^\gamma \) to greater than \( 0.6n \) “instantaneously”

\[ C_{\text{max}} \text{ jumps from sublinear } n^\gamma \]

to \( \geq 0.5n \) in \( n^\beta \) edges, with \( \beta, \gamma < 1 \).

Nontrivial Scaling behaviors
\( \gamma + 1.2\beta = 1.3 \) for \( A \in [0.1, 0.6] \)

Achlioptas, D’Souza, Spencer, *Science*, **323** (5920), 2009
A Hybrid Transition! Diverging correlation length, I

The second largest component, $C_2$

$$C_2 \sim (t_c - t)^{-\gamma}, \text{ with } \gamma \approx 1.15$$

(No simple corrections to scaling yield mean field $\gamma = 1$)
• “Susceptibility”, the expected size of the component to which an arbitrarily chosen vertex belongs,

\[ \chi = \frac{1}{n} \sum_{v \in V(G)} |C(v)| = \frac{1}{n} \sum_{\text{components}} |C_i|^2. \]

\[ \chi \sim (t_c - t)^{-\gamma}, \text{ with } \gamma \approx 1.17 \]
Explosive percolation now observed in ...

  “Explosive Growth in Biased Dynamic Percolation on Two-Dimensional Regular Lattice Networks”

  “Percolation Transitions in Scale-Free Networks under the Achlioptas Process”
  (Chung-Lu weighted node power law growth model)
  \[ p_c > 0 \text{ for } \gamma > 2.3 \text{ or } 2.4 \text{ and discontinuous.} \]

  “Explosive percolation in scale-free networks”
  (Configuration model power law)
  \[ p_c > 0 \text{ for } \gamma > 2.2, \text{ discontinuous for } \gamma > 3. \]

  “Construction and Analysis of Random Networks with Explosive Percolation”

  “Cluster aggregation model for discontinuous percolation transition”

- Rozenfeld, Gallos, Makse; arxiv:0911.4082
  “Explosive Percolation in the Human Protein Homology Network”
Beyond “Product Rule”

- “Sum rule” also works, but delay is smaller.
- In general any rule that keeps components similar in size in subcritical regime should be explosive:
  - “Powder Keg” of Friedman and Landsberg *PRL* (2009).
  - Starting ER from proper initial state; Cho, Khang, Kim *PRE* (2010).
Rigorous techniques only for **bounded size** (thus far)

- Bounded size rules
  (treat all components of size $\geq K$ as the same)

- Assume cluster aggregation models (two distinct clusters merged with each edge addition).

- “Birth Control for Giants”,
  J. Spencer, N. Wormald, *Combinatorica* **27**(5), 2007:
  - Differential equation for evolution
    (rigorous proof that error term is small in subcritical regime)
  - Conjecture that all bounded size rules have continuous PT’s
Local Cluster Aggregation Models with Explosive Percolation


- “Adjacent Edge (AE)” — 2 candidate edges share a common vertex.
- “Triangle Rule” (TR) — choose 3 vertices at random, 3 candidate edges.
Locality: more physical and simplifies analytic treatment

Adjacent Edge Rule (3 components) vs Product Rule (4 components)

(TR also depends only on 3 components)
Evolution equations for the bounded size AE rule:

- Let $x_i$ denote fraction of nodes in components of size $i$, for $1 \leq i \leq K$.

- Let $S_i = \sum_{j=i}^{\infty} x_j$
  (the weight in the tail starting at size $i$)
  ($S_{K+1} = 1 - \sum_{j=1}^{K} x_j$ is useful)

- Probability first node is in component of size $i$ is $x_i$.

- Probability the smaller of the two additional components has size $j$ is $S_j^2 - S_{j+1}^2$.

- Expected evolution of $x_i$'s:

$$\frac{dx_i}{dt} = -i x_i - i(S_i^2 - S_{i+1}^2) + i \sum_{j+k=i} x_j(S_k^2 - S_{k+1}^2)$$
“Susceptibility”: \[ W = \sum_{i=1}^{\infty} ix_i = \sum_{i=1}^{\infty} i^2 n_i \]

(Expected size of component to which arbitrary vertex belongs)

The evolution of \( W \):

\[
\frac{dW}{dt} = \sum_{j=1}^{K} \sum_{k=1}^{K} 2jkx_j(s_k^2 - s_{k+1}^2) + \sum_{j=1}^{K} 2jW^*x_j s_{K+1} \\
+ \sum_{k=1}^{K} 2kW^*(s_k^2 - s_{k+1}^2) + 2(W^*)^2 s_{K+1}.
\]

• Where \( W^* = W - \sum_{i=1}^{K} ix_i \)
  (contributions from comps of size greater than \( K \)).

• \( W \to \infty \) at the phase transition. Simple Euler’s method numerics yields \( t_c = 0.796 \) (with \( K = 600 \)).
Direct simulation of the AE graph evolution process

- Find agreement to three digits, \( t_c = 0.796 \).

- \( \Delta \) sublinear in \( n \), where \( t_0 \) is last time \( C_1 \leq n^\gamma \), and \( t_1 \) first time \( C_1 \geq An \). (Denote this by \( \Delta(\gamma, A) \))

(For AE \( \Delta(0.5, 0.2) \) sublinear. For PR \( \Delta(0.5, 0.6) \) sublinear.)
AE, PR, TR are Hybrid Transitions!
(Discontinuous change, but scaling behavior)

- Component density $n_i \sim i^{-\tau}$
- $W \sim |t - t_c|^{-\alpha}$
- $C_2 \sim |t - t_c|^{-\mu}$

<table>
<thead>
<tr>
<th></th>
<th>PR</th>
<th>AE</th>
<th>TR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_c$</td>
<td>0.888</td>
<td>0.796</td>
<td>0.848</td>
</tr>
<tr>
<td>$\tau$</td>
<td>2.1</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1.17</td>
<td>1.13</td>
<td>1.13</td>
</tr>
<tr>
<td>$\mu$</td>
<td>1.17</td>
<td>1.13</td>
<td>1.13</td>
</tr>
</tbody>
</table>
Other Hybrid Transitions

- **k-sat** (constraint satisfaction) for $k \geq 3$, Infinite dimensional

- **Jamming in models of granular materials**
  Finite dimensional / spatial constraints


- **Spin glasses** glassy systems, slow relaxation time

Mixed transitions: geometry, disorder, computation

- Critical slowing down and computational complexity?
- *Nature* 1999: 2-sat continuous $\in P$, 3-sat discontinuous $\in NP$
- Hard instances, applications to Cryptography?

- Aspen Center for Physics: “*Complexity, Disorder, and Algorithms*”
  Organizers: S. Coppersmith, A. Middleton, J. Machta, C. Moore

  Organizers: P. Diaconis, D. Fisher, C. Moore, C. Radin

- Mathematical Sciences Research Institute (MSRI)
  “*Probability, Algorithms and Statistical Physics*”
  Jan 3 - May 15, 2005.
A collection of interacting networks

Networks:
- Transportation Networks/Power grid (distribution/collection networks)
- Biological networks - protein interaction - genetic regulation - drug design
- Computer networks
- Social networks - Immunology - Information - Commerce
Modular Erdős-Rényi

- Divide nodes initially into two groups (A and B):

- Add internal $a-a$ edges with rate $\lambda$.
- Add internal $b-b$ edges with rate $\lambda/r_1$, with $r_1 > 1$.
- Add intra-group $a-b$ edges with rate $\lambda/r_2$, with $r_2 > 1$, $r_2 \neq r_1$.

What happens? (Anything different?)
Percolation on interacting networks, using random graph models

System of two networks
Connectivity for an individual node

- Probability distribution nodes in network \( a \): \( p^a_{k_a k_b} \)
- For the system: \( \{ p^a_{k_a k_b}, p^b_{k_a k_b} \} \)
- Build generating function formalism for interacting networks.
Generating Functions – Distribution of component sizes:
(Extending Newman, Strogatz, Watts *PRE* 64, (2001))

Three step process:

1. G.F. for connectivity of a node connected to a random edge, \( G_{ab}(x_a, x_b) \).
2. G.F. for the size of the component to which that node belongs, \( H_{ab}(x_a, x_b) \).
3. G.F. for the size of the component to which an arbitrary node belongs, \( H_a(x_a, x_b) \).

Moments of GFs provide information, e.g., the expected number of \( a \)-nodes in the component of an arbitrary \( a \)-node:

\[
\langle s_a \rangle_a = \left. \frac{\partial}{\partial x_a} H_a(x_a, x_b) \right|_{x_a=x_b=1}
\]
Distributions for modular Erdős-Rényi

- \( p_{k_a k_b}^a = p_{k_a}^a p_{k_b}^a \) (uncorrelated)

- Independent Poisson distributions with related means:
  \[
  \overline{k}_{bb} = \frac{1}{r_1} \overline{k}_{aa},
  \overline{k}_{ab} = \overline{k}_{ba} = \frac{1}{2r_2} \overline{k}_{aa}.
  \]

\[
\begin{align*}
  p_{k_a k_b}^a &= \frac{\overline{k}_{aa}^{k_a} e^{-\overline{k}_{aa}}}{k_a!} \cdot \frac{\overline{k}_{ab}^{k_b} e^{-\overline{k}_{ab}}}{k_b!} \\
  p_{k_a k_b}^b &= \frac{\overline{k}_{bb}^{k_b} e^{-\overline{k}_{bb}}}{k_b!} \cdot \frac{\overline{k}_{ba}^{k_a} e^{-\overline{k}_{ba}}}{k_a!}
\end{align*}
\]

\[
\langle s_a \rangle_a = 1 + \frac{\overline{k}_{aa} - \overline{k}_{aa} \overline{k}_{bb} + \overline{k}_{ab} \overline{k}_{ba}}{(1 - \overline{k}_{aa})(1 - \overline{k}_{bb}) - \overline{k}_{ab} \overline{k}_{ba}}
\]

(Bollobas, Janson, Riordan RSA, 2007
Ostilli, Mendes J Stat Mech 2009)
Wiring which respects group structures percolates earlier!

Fraction of nodes $\pm k$

$S^{L1}$

$S^{L2}$

$S^{ER}$

“Cooperative enhancement”

aa rate $\lambda$

bb rate $\lambda/r_1$, w/ $r_1 = 2$

ab rate $\lambda/r_2$, w/ $r_2 = 6$
1. **Two interacting trunc PLs**

\[ p_k \sim k^{-\tau} \exp(-k/\kappa) \]

- \( \tau_a = 2.8, \kappa_a = 100 \)
- \( \tau_b = 2.0, \kappa_b \text{ varied} \)

- \( a) \) \( \kappa_a b = p_{k_a}^b \text{ Poisson } \bar{k}_{ab} = 0.4 \)

- \( b) \) \( \kappa_b \text{ varied} \)

- \( ab) \) \( p_{k_b}^a = p_{k_a}^b \text{ Poisson } \bar{k}_{ab} = 0.4 \)

- Approximates two loosely coupled human contact networks.

- Giant component is larger for \( L2 \).
Other distributions, cont.

2. **Trunc PL and Poisson**
   
   with \( n_a = 4n_b \)
   
   a) Poisson, \( \bar{k}_{aa} = 0.5 \)
   
   b) \( \tau_b = 2.5, \kappa_b \) varied
   
   ab) \( p_{k_b}^a \) Poisson \( \bar{k}_{ab} = 0.2 \)

- Approximates critical infrastructure:
  
  – power grid nodes having narrow degree distribution.
  
  – Internet routers have broad degree distribution.
Statistical signatures of interacting networks: (How and when do they differ from random or each-other?)

- **Socio-technical congruence** (e.g., Open Source Software)

- **Searching for biomarkers of disease:**
  - Protein-gene interactions
  - Protein-protein interactions
  - Proteome
  - Metabolism
  - Bio-chemical reactions
  - Cellular networks:
Explosive percolation in random graphs – Conclusions

- Controlling phase transitions with choice
  - Delay or enhance
  - Changing speed of onset
  - Changing universality classes
  - Including locality

- Network interactions
  - Can change the onset of phase transitions
  - Modular treatments can percolate sooner (or later, depending on specifics of degree distributions involved)

- Random graphs and real-world networks
  - Can the differences serve as diagnostic tools?