Cascades on interdependent networks

Raissa M. D’Souza
University of California, Davis
Complexity Sciences Center
External Professor, Santa Fe Institute
A collection of interacting, dynamic networks form the core of modern society.

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

- E-commerce → WWW → Internet → Power grid → River networks.
- Biological virus → Social contact network → Transportation nets → Communication nets → Power grid → River networks.
Moving to systems of interdependent networks

What are the simplest, useful, abstracted models?

- What are the emergent new properties?
  - Host-pathogen interactions
  - Phase transition thresholds

- What features confer resilience in one network while introducing vulnerabilities in others?

- How do demands in one system shape the performance of the others? (e.g., demand informed by social patterns of communication)

- How do constraints on one system manifest in others? (e.g., River networks shape placement of power plants)

- Coupling of scales across space and time / co-evolution.
Configuration model for interacting networks


System of two networks

Connectivity for an individual node

- Degree distribution for nodes in network $a$: $p_{k_ak_b}^a$
- For the system: $\{p_{k_ak_b}^a, p_{k_ak_b}^b\}$
- Generating functions to calculate properties of the ensemble of such networks.
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?)
Wiring which respects group structures percolates earlier!

(Also tradeoffs between sparser and denser subnetworks.)

- Probability distribution for node degrees: \( \{ p^a_{k_ak_b}, p^b_{k_ak_b} \} \)
- Generating functions to calculate properties of the ensemble of such networks.
Consider two coupled random graphs.

Nodes fail (removed either in a targeted or random manner).

Following an **iterative removal process**, small failures can lead to massive cascades of failure of the networks themselves.

**Surprising**: What confers resilience to individual network (broad-scale degree distribution) may be a weakness for randomly coupled networks.
Single networks – broad scale degree distribution distribution.

Social contacts
Szendrői and Csányi

Airport traffic
Bounova 2009

Approximated as power law $P_k \propto k^{-\gamma}$
$P_k \sim k^{-\gamma}$, the first two moments

(Note: $\gamma > 1$ required for $\sum_k P_k = 1$)

- First moment (Mean degree):

$$\langle k \rangle = \sum_{k=1}^{\infty} kp_k \approx \int_{k=1}^{\infty} kp_k dk$$

Diverges (i.e., $\langle k \rangle \rightarrow \infty$) if $\gamma \leq 2$.

- Second moment:

$$\langle k^2 \rangle = \sum_{k=1}^{\infty} k^2 p_k \approx \int_{k=1}^{\infty} k^2 p_k dk$$

Diverges (i.e., $\langle k^2 \rangle \rightarrow \infty$) if $\gamma \leq 3$.

- Many results follow for $2 < \gamma < 3$ since $\langle k \rangle / \langle k^2 \rangle \rightarrow 0$
Consequences of \( p(k) \sim k^{-\gamma} \) for networks

- Most nodes are leaves (degree 1): Network connectivity very robust to random node removal.
- High degree nodes are hubs: Network connectivity very fragile to targeted node removal.

Epidemic spreading on the network (contact process):

If \( 2 < \gamma < 3 \), then \( \langle k \rangle / \langle k^2 \rangle \rightarrow 0 \) and \( \lim_{n \rightarrow 0} \) epidemic threshold \( \rightarrow 0 \).

(Buldyrev et al find broad scale more fragile for their particular cascade dynamics)
Dynamical processes on interdependent networks

Motivation: interconnected power grids


Power grid: a collection of interdependent grids. (Interconnections built originally for emergencies.)

Blackouts cascade from one grid to another (in a non-local manner).

Building more interconnections (Fig: planned wind transmission).

Increasingly distributed

Source: NPR
What is the effect of interdependence on cascades?

It is thought power grids organize to a "critical" state – power law distribution of black out sizes – maximize profits while fearing large cascades.

Source: NPR
Sandpile models: “Self-organized criticality”

- Drop grains of sand ("load") randomly on nodes.
- Each node has a threshold for sand.
- Load $> \text{threshold} \Rightarrow \text{node topples} = \text{sheds sand to neighbors.}$
- These neighbors may topple. And their neighbors. And so on.
- Cascades of load/stress on a system.

The classic Bak-Tang-Wiesenfeld sandpile model:
(Neuronal avalanches, banking cascades, earthquakes, landslides, forest fires, blackouts...)

- Finite square lattice in $\mathbb{Z}^2$
- Thresholds 4
- Open boundaries prevent inundation

Avalanche size follows power law distribution $P(s) \sim s^{-3/2}$
Sandpile model on arbitrary networks:

- Thresholds = degrees (shed one grain per neighbor)
- Boundaries: shedded sand are deleted independently with probability $f (\approx 10/N)$
- Mean-field behavior ($P(s) \propto s^{-3/2}$) robust.
  (Goh et al. PRL 03, Phys. A 2004/2005, PRE 2005. PLRGs with $2 < \gamma < 3$ not mean-field.)

Sandpiles on interacting networks:

- Sparse connections between random graphs.
- Configuration model with multi-type degree distribution.
Sparsely coupled networks

Two-type network: $a$ and $b$.

Degree distributions: $p_a(k_a, k_b), p_b(k_a, k_b)$

$$p_a(k_a, k_b) = \text{fraction of } a\text{-nodes with } k_a, k_b \text{ neighbors in } a, b.$$ 

Configuration model: create degree sequences until valid (even total intra-degree, same number of inter-edge stubs), then connect edge stubs at random.
Measures of avalanche size

- **Topplings:**
  Drop a grain of sand. How many nodes eventually topple?

  **Avalanche size** distributions: \( s_a(t_a, t_b), s_b(t_a, t_b) \)

  e.g., \( s_a(t_a, t_b) = \text{chance an avalanche begun in } a \text{ topples } t_a \text{ many } a\text{-nodes, } t_b \text{ many } b\text{-nodes.} \)

  To study this, we need a more basic distribution...

- **Sheddings:**
  Drop a grain of sand. How many grains are eventually shed from one network to another?

  **Shedding size** distributions: \( \rho_{od}(r_{aa}, r_{ab}, r_{ba}, r_{bb}) \)

  = \text{chance a grain shed from network } o \text{ to } d \text{ eventually causes } r_{aa}, r_{ab}, r_{ba}, r_{bb} \text{ many grains to be shed from } a \rightarrow a, a \rightarrow b, b \rightarrow a, b \rightarrow b \)

*Approximate shedding and toppling as multi-type branching processes.*
Cascades in networks $\approx$ branching processes if they’re tree-like.

Power grids are fairly tree-like:

<table>
<thead>
<tr>
<th></th>
<th>clustering coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power grid in SE USA</td>
<td>0.01</td>
</tr>
<tr>
<td>Similar Erdős-Rényi graph</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Sandpile cascades on interacting networks $\approx$ a multitype branching process.
Overview of the calculations

From degree distribution to avalanche size distribution:

**Input**: degree distributions $p_a(k_a, k_b), p_b(k_a, k_b)$

\[ \downarrow \text{compute} \]

*sheddling* branching distributions $q_{aa}, q_{ab}, q_{ba}, q_{bb}$

\[ \downarrow \text{compute} \]

*toppling* branching distributions $u_a, u_b$

\[ \downarrow \text{plug in} \]

*toppling* branching generating functions $U_a, U_b$

\[ \downarrow \text{plug in} \]

equations for avalanche size generating functions $S_a, S_b$

\[ \downarrow \text{solve numerically, asymptotically} \]

**Output**: avalanche size distributions $s_a, s_b$
Example:

\[ q_{ab}(r_{ba}, r_{bb}) := \text{the branch (children) distribution for an } \text{ab-shedding}. \]

Probability a single grain shed from \( a \) to \( b \) results in \( r_{ba} \) \text{-sheddings and } \( r_{bb} \) \text{-sheddings.}
Shedding branch distributions \(q_{od}\)

The crux of the derivation

\[
q_{od}(r_{da}, r_{db}) := \text{chance a grain of sand shed from network } o \text{ to } d \\
\text{topples that node, sending } r_{da}, r_{db} \text{ many grains to networks } a, b.
\]

\[
q_{od}(r_{da}, r_{db}) = \frac{r_{do}p_d(r_{da}, r_{db})}{\langle k_{do} \rangle} \frac{1}{r_{da} + r_{db}} \\
\text{for } r_{da} + r_{db} > 0.
\]

- I: chance the grain lands on a node with degree \(p_d(r_{da}, r_{db})\) (Edge following: \(r_{do}\) edges leading from network \(o\).)

- II: empirically, sand on nodes is \(\sim\) Uniform\(\{0, ..., k - 1\}\)

- Chance of no children = \(q_{od}(0, 0) := 1 - \sum_{r_{da}+r_{db}>0} q_{od}(r_{da}, r_{db})\) (Probability a neighbor of any degree sheds, properly weighted.)

- Chance at least one child = \(1 - q_{od}(0, 0)\).
I. Edge following probability: single network

- Degree distribution, $P_k$, with G.F. $G_0(x) = \sum_k P_k x^k$.
- Probability of following a random edge to a node of degree $k$: $q_k = kP_k/\sum_k kP_k$, with G.F. $G_1(x) = \sum_k q_k x^k$.
- (Edge following: “Contact immunization” strategy of CDC.)
- Generating function “self consistency” construction.

$H_1(x)$: G.F. for dist in comp size following random edge

\[
H_1(x) = xq_0 + xq_1 H_1(x) + xq_2 [H_1(x)]^2 + xq_3 [H_1(x)]^3 \cdots \\
= xG_1(H_1(x))
\]

(c.f. Newman, Strogatz, Watts *PRE* 2001.)
II. Revisiting the “1/k” assumption

Pierre-André Noël, C. Brummitt, R. D’Souza
in progress

A node that just toppled is actually less likely to topple on the next time step.
(prob zero sand $\neq 1/k$)
Toppling branch distributions  \( u_a, u_b \)
shedding branch distributions  \( q_{od} \sim \) toppling branch distributions  \( u_a, u_b \)

**Key:** a node topples iff it sheds at least one grain of sand.

Probability an  \( o \) to  \( d \) shedding leads to at least one other shedding:  \( 1 - q_{od}(0, 0) \). Probability a single shedding from an  \( a \)-node yields  \( t_a, t_b \) topplings:

\[
u_a(t_a, t_b) = \sum_{k_a=t_a, k_b=t_b}^{\infty} p_a(k_a, k_b) \text{Binomial}[t_a; k_a, 1 - q_{aa}(0, 0)] \cdot \text{Binomial}[t_b; k_b, 1 - q_{ab}(0, 0)].
\]

(e.g.,  \( k_a \) neighbors,  \( t_a \) of them topple, each topples with prob  \( 1 - q_{aa}(0, 0) \).)

Associated generating functions:  \( U_a(\tau_a, \tau_b), U_b(\tau_a, \tau_b) \).
Summary of distributions and their generating functions

<table>
<thead>
<tr>
<th>degree</th>
<th>distribution</th>
<th>generating function</th>
</tr>
</thead>
<tbody>
<tr>
<td>shedding branch</td>
<td>$p_a(k_a, k_b), p_b(k_a, k_b)$</td>
<td>$G_a(\omega_a, \omega_b), G_b(\omega_a, \omega_b)$</td>
</tr>
<tr>
<td>toppling branch</td>
<td>$q_{od}(r_{da}, r_{db})$</td>
<td></td>
</tr>
<tr>
<td>toppling size</td>
<td>$u_a(t_a, t_b), u_b(t_a, t_b)$</td>
<td>$U_a(\tau_a, \tau_b), U_b(\tau_a, \tau_b)$</td>
</tr>
<tr>
<td></td>
<td>$s_a(t_a, t_b), s_b(t_a, t_b)$</td>
<td>$S_a(\tau_a, \tau_b), S_b(\tau_a, \tau_b)$</td>
</tr>
</tbody>
</table>

Self-consistency equations:

$$S_a = \tau_a U_a(S_a, S_b), \quad (1)$$
$$S_b = \tau_b U_b(S_a, S_b). \quad (2)$$

Want to solve (1), (2) for $S_a(\tau_a, \tau_b), S_b(\tau_a, \tau_b)$.

Coefficients of $S_a, S_b =$ avalanche size distributions $s_a, s_b$.

In practice, Eqs. (1), (2) are transcendental and difficult to invert.
Numerically solving $\vec{S}(\vec{\tau}) = \vec{\tau} \cdot \vec{U}(\vec{S}(\vec{\tau}))$

Methods for computing $s_a, s_b$ for small avalanche size:

**Method 1:** Iterate starting from $S_a = S_b = 1$; expand.

**Method 2:** Iterate symbolically; use Cauchy’s integration formula

$$s_a(t_a, t_b) = \frac{1}{(2\pi i)^2} \int \int_D \frac{S_a(\tau_a, \tau_b)}{\tau_a^{t_a+1} \tau_b^{t_b+1}} d\tau_a d\tau_b,$$

where $D \subset \mathbb{C}^2$ encloses the origin and no poles of $S_a$.

**Method 3:** Multidimensional Lagrange inversion (IJ Good 1960):

$$S_a = \sum_{m_a, m_b = 0}^{\infty} \frac{\tau_a^{m_a} \tau_b^{m_b}}{m_a! m_b!} \left[ h(\vec{\kappa}) U_a(\vec{\kappa})^m a U_b(\vec{\kappa})^m b \left| \frac{\delta_{\nu} - \frac{\kappa_{\mu}}{U_{\mu}} \frac{\partial U_{\mu}}{\partial \kappa_{\mu}}}{\kappa_{\mu}} \right|_{\vec{\kappa} = 0} \right],$$

if the types $\mu, \nu \in \{a, b\}$ have a positive chance of no children.

- Unfortunately for large avalanches need to use simulation.
  (Asymptotic approximations used for isolated networks do not apply.)
Two geographically nearby power grids in the southeastern US.

<table>
<thead>
<tr>
<th></th>
<th>Grid c</th>
<th>Grid d</th>
</tr>
</thead>
<tbody>
<tr>
<td># nodes</td>
<td>439</td>
<td>504</td>
</tr>
<tr>
<td>$\langle k_{int} \rangle$</td>
<td>2.4</td>
<td>2.9</td>
</tr>
<tr>
<td>$\langle k_{ext} \rangle$</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>clustering</td>
<td>0.01</td>
<td>0.08</td>
</tr>
</tbody>
</table>

8 links between these two distinct grids. Different average internal degree $\langle k_{int} \rangle$. Long paths. (Low clustering – approximately locally tree-like.)
Two random $z_a$, $z_b$-regular graphs with “Bernoulli coupling”: each node gets an external link independently with probability $p$. These $\approx$ power grids.

\[
\mathcal{U}_a(\tau_a, \tau_b) = \frac{(p - p\tau_a + (z_a + 1)(\tau_a + z_a - 1))^{z_a}(1 + p(\tau_b - 1) + z_b)}{(z_a + 1)^{z_a}z_a^{z_a}(z_b + 1)}
\]
Matching theory and simulation (for small’ish avalanches)

Plot marginalized avalanche size distributions

\[ s_a(t_a) \equiv \sum_{t_b \geq 0} s_a(t_a, t_b), \quad s_a(t_b) \equiv \sum_{t_a \geq 0} s_a(t_a, t_b), \quad \text{etc.} \]

in simulations, branching process.

Regular(3)-Bernoulli(p)-Regular(10)  Power grids c, d.
Main findings: For an individual network, optimal $p^*$

- **(Blue curve)** Initially increasing $p$ decreases the largest cascades started in that network (second network is reservoir for load).
- **(Red curve)** Increasing $p$ increases the largest cascades inflicted from the second network (two reasons: new channels and greater capacity).
- **(Gold curve)** Neglecting the origin of the cascade, the effects balance at a stable critical point, $p^* \approx 0.1$. (Reduced by 75% from $p=0.001$ to $p=0.1$)
Main findings: Individual network, “Yellowstone effect”

Supressing largest cascades amplifies small and intermediate ones! (Supressing smallest amplifies largest (Yellowstone and Power Grids*))

\[
\begin{align*}
\text{Pr}(1 \leq T_a \leq 50) & \quad \text{small cascades} \\
\text{Pr}(50 \leq T_a \leq 99) & \\
\text{Pr}(250 \leq T_a \leq 299) & \\
\text{Pr}(350 \leq T_a \leq 399) & \quad \text{large cascades}
\end{align*}
\]

- To suppress smallest, isolation \( p = 0 \).
- To suppress intermediate (10% of system size) either \( p = 0 \) or \( p = 1 \).
- To suppress cascades > 25% of system size then \( p = p^* \approx 0.11 \).

*Dobson I, Carreras BA, Lynch VE, Newman DE Chaos, (2007).*
Main findings: System as a whole

More interconnections fuel larger system-wide cascades.

- Each new interconnection adds capacity and load to the system (Here capacity is a node’s degree, interconnections increase degree)

- Test this on coupled random-regular graphs by rewiring internal edges to be spanning edges (increase interconnectivity with out increasing degree). No increase in the largest cascades.

- Inflicted cascades (Red curve) increase mostly due to increased capacity.

- So an individual operator adding edges to achieve $p^*$ may inadvertently cause larger global cascades.
Larger cascades from increased interconections:
A warning sign?

- Financial markets
- Energy transmission systems

Unless the coupled grids are identical, only one will be able to achieve it's $p^*$. 

- Coupled $z_a \neq z_b$ regular random graphs (branching process and simulation).

\[
\frac{\langle s_a \rangle b}{\langle s_b \rangle a} = \frac{1 + z_a}{1 + z_b}
\]

If $z_b > z_a$ inflicted cascades from $b$ to $a$ larger than those from $a$ to $b$.

(An arm’s race for capacity?)
Summary: Sandpile cascades on interacting networks

- Some interconnectivity can be *beneficial*, but too much is *detrimental*. Stable optimal levels are possible.

- From perspective of isolated network, seek optimal interconnectivity $p^*$. 

- This *equilibrium will be frustrated* if the two networks differ in their load or propensity to cascade.

- Tuning $p$ to *suppress* large cascades *amplifies* to occurrence of small ones. (Likewise, suppressing small, amplifies large.)

- Additional capacity and overall load from new interconnections *fuels larger cascades* in the system as a whole.

- What might be good for an individual operator (adding edges to achieve $p^*$), may be bad for society.
Possible extensions – Real power grids

- Expand multi-type processes to encode for different types of nodes (buses, transformers, generators)
- Linearized power flow equations – cascades in real power grids are non-local: e.g. fig: 3 to 4, 7 to 8
- Game theoretic/economic consideration (we assume adding connections is cost-free)

(1996 Western blackout NERC report)

(Power grids as “critical” – Balancing profit and fear of outages)
Possible extensions

Teams and social networks
- Tasks (sand) arriving on people (nodes)
- Each person has a capacity for tasks: sheds once overloaded
- Coupling to a second social network (team) can reduce large cascades

Amplifying cascades
- Encourage adoption of new products
- Snowball sampling

Airline networks
- Different carriers accepting load (bumped passengers)
Other types of cascades, not just than sandpiles

- Watt’s threshold model: “topple” is some fraction $\phi$ of your neighbors have “toppled” (rather than “toppling”, Watt’s think of cascades in adopting a new product).
  - Harder to “topple” nodes of high degree.

- Kleinberg: rather than thresholds, diminishing returns (concave / sub-modular utility)

Note Author Summary for high-level overview.