Mixed, correlated random networks

Last updated: 2021/10/07, 17:43:00 EDT

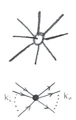
Principles of Complex Systems, Vols. 1 & 2
CSYS/MATH 300 and 303, 2021–2022 | @pocsvox

Prof. Peter Sheridan Dodds | @peterdodds

Computational Story Lab | Vermont Complex Systems Center
Vermont Advanced Computing Core | University of Vermont

Outline

Random directed networks:



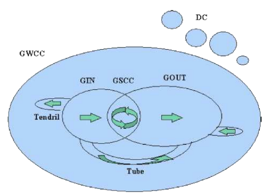
- So far, we've largely studied networks with undirected, unweighted edges.
- Now consider directed, unweighted edges.
- Nodes have k_i and k_o incoming and outgoing edges, otherwise random.
- Network defined by joint in- and out-degree distribution: P_{k_i,k_o}
- Normalization: $\sum_{k_i=0}^{\infty}\sum_{k_o=0}^{\infty} P_{k_i,k_o} = 1$
- Marginal in-degree and out-degree distributions:

$$P_{k_i} = \sum_{k_o=0}^{\infty} P_{k_i,k_o} \text{ and } P_{k_o} = \sum_{k_i=0}^{\infty} P_{k_i,k_o}$$

- Required balance:

$$\langle k_i \rangle = \sum_{k_i=0}^{\infty}\sum_{k_o=0}^{\infty} k_i P_{k_i,k_o} = \sum_{k_i=0}^{\infty}\sum_{k_o=0}^{\infty} k_o P_{k_i,k_o} = \langle k_o \rangle$$

Directed network structure:

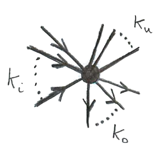


From Boguñá and Serano. [1]

- GWCC = Giant Weakly Connected Component (directions removed);
- GIN = Giant In-Component;
- GOUT = Giant Out-Component;
- GSCC = Giant Strongly Connected Component;
- DC = Disconnected Components (finite).
- When moving through a family of increasingly connected directed random networks, GWCC usually appears before GIN, GOUT, and GSCC which tend to appear together. [4, 1]

Important observation:

- Directed and undirected random networks are separate families ...
- ...and analyses are also disjoint.
- Need to examine a larger family of random networks with mixed directed and undirected edges.



- Consider nodes with three types of edges:
  1. k_u undirected edges,
  2. k_i incoming directed edges,
  3. k_o outgoing directed edges.
- Define a node by generalized degree:

$$\vec{k} = [\, k_u \;\; k_i \;\; k_o \,]^{T}.$$

- Joint degree distribution:

$$P_{\vec{k}} \text{ where } \vec{k} = [\, k_u \;\; k_i \;\; k_o \,]^{T}.$$

- As for directed networks, require in- and out-degree averages to match up:

$$\langle k_i \rangle = \sum_{k_u=0}^{\infty}\sum_{k_i=0}^{\infty}\sum_{k_o=0}^{\infty} k_i P_{\vec{k}} = \sum_{k_u=0}^{\infty}\sum_{k_i=0}^{\infty}\sum_{k_o=0}^{\infty} k_o P_{\vec{k}} = \langle k_o \rangle$$

- Otherwise, no other restrictions and connections are random.
- Directed and undirected random networks are disjoint subfamilies:

$$\text{Undirected: } P_{\vec{k}} = P_{k_u}\delta_{k_i,0}\delta_{k_o,0},$$

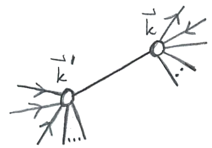
$$\text{Directed: } P_{\vec{k}} = \delta_{k_u,0}P_{k_i,k_o}.$$

Correlations:

- Now add correlations (two point or Markovian) ☐:
  1. $P^{(u)}(\vec{k}\,|\,\vec{k}')$ = probability that an undirected edge leaving a degree $\vec{k}'$ nodes arrives at a degree $\vec{k}$ node.
  2. $P^{(i)}(\vec{k}\,|\,\vec{k}')$ = probability that an edge leaving a degree $\vec{k}'$ nodes arrives at a degree $\vec{k}$ node is an in-directed edge relative to the destination node.
  3. $P^{(o)}(\vec{k}\,|\,\vec{k}')$ = probability that an edge leaving a degree $\vec{k}'$ nodes arrives at a degree $\vec{k}$ node is an out-directed edge relative to the destination node.
- Now require more refined (detailed) balance.
- Conditional probabilities cannot be arbitrary.
  1. $P^{(u)}(\vec{k}\,|\,\vec{k}')$ must be related to $P^{(u)}(\vec{k}'\,|\,\vec{k})$.
  2. $P^{(o)}(\vec{k}\,|\,\vec{k}')$ and $P^{(i)}(\vec{k}\,|\,\vec{k}')$ must be connected.

Correlations—Undirected edge balance:

- Randomly choose an edge, and randomly choose one end.
- Say we find a degree $\vec{k}$ node at this end, and a degree $\vec{k}'$ node at the other end.
- Define probability this happens as $P^{(u)}(\vec{k},\vec{k}')$.
- Observe we must have $P^{(u)}(\vec{k},\vec{k}') = P^{(u)}(\vec{k}',\vec{k})$.



- Conditional probability connection:

$$P^{(u)}(\vec{k},\vec{k}') = P^{(u)}(\vec{k}\,|\,\vec{k}')\frac{k_u' P(\vec{k}')}{\langle k_u' \rangle}$$

$$P^{(u)}(\vec{k}',\vec{k}) = P^{(u)}(\vec{k}'\,|\,\vec{k})\frac{k_u P(\vec{k})}{\langle k_u \rangle}.$$

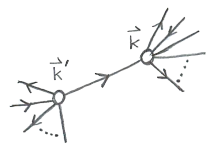
Correlations—Directed edge balance:

- The quantities

$$\frac{k_o P(\vec{k})}{\langle k_o \rangle} \text{ and } \frac{k_i P(\vec{k})}{\langle k_i \rangle}$$

give the probabilities that in starting at a random end of a randomly selected edge, we begin at a degree $\vec{k}$ node and then find ourselves travelling:

1. along an outgoing edge, or
2. against the direction of an incoming edge.



- We therefore have

$$P^{(dir)}(\vec{k},\vec{k}') = P^{(i)}(\vec{k}\,|\,\vec{k}')\frac{k_o' P(\vec{k}')}{\langle k_o' \rangle} = P^{(o)}(\vec{k}'\,|\,\vec{k})\frac{k_i P(\vec{k})}{\langle k_i \rangle}.$$

- Note that $P^{(dir)}(\vec{k},\vec{k}')$ and $P^{(dir)}(\vec{k}',\vec{k})$ are in general not related if $\vec{k} \neq \vec{k}'$.

Global spreading condition: [2]

When are cascades possible?:

- Consider uncorrelated mixed networks first.
- Recall our first result for undirected random networks, that edge gain ratio must exceed 1:

$$\mathbf{R} = \sum_{k_u=0}^{\infty} \frac{k_u P_{k_u}}{\langle k_u \rangle} \bullet (k_u - 1) \bullet B_{k_u,1} > 1.$$

- Similar form for purely directed networks:

$$\mathbf{R} = \sum_{k_i=0}^{\infty} \sum_{k_o=0}^{\infty} \frac{k_i P_{k_i,k_o}}{\langle k_i \rangle} \bullet k_o \bullet B_{k_i,1} > 1.$$

- Both are composed of (1) probability of connection to a node of a given type; (2) number of newly infected edges if successful; and (3) probability of infection.

Global spreading condition:

Local growth equation:

- Define number of infected edges leading to nodes a distance d away from the original seed as $f(d)$.
- Infected edge growth equation:

$$f(d+1) = \mathbf{R} f(d).$$

- Applies for discrete time and continuous time contagion processes.
- Now see $B_{k_u,1}$ is the probability that an infected edge eventually infects a node.
- Also allows for recovery of nodes (SIR).

Global spreading condition:

Mixed, uncorrelated random netwoks:

- Now have two types of edges spreading infection: directed and undirected.
- Gain ratio now more complicated:
  1. Infected directed edges can lead to infected directed or undirected edges.
  2. Infected undirected edges can lead to infected directed or undirected edges.
- Define $f^{(u)}(d)$ and $f^{(o)}(d)$ as the expected number of infected undirected and directed edges leading to nodes a distance d from seed.

- Gain ratio now has a matrix form:

$$\begin{bmatrix} f^{(u)}(d+1) \\ f^{(o)}(d+1) \end{bmatrix} = \mathbf{R} \begin{bmatrix} f^{(u)}(d) \\ f^{(o)}(d) \end{bmatrix}$$

- Two separate gain equations:

$$f^{(u)}(d+1) = \sum_{\vec{k}} \left[\frac{k_u P_{\vec{k}}}{\langle k_u \rangle} \bullet (k_u - 1) \bullet B_{k_u+k_i,1} f^{(u)}(d) + \frac{k_i P_{\vec{k}}}{\langle k_i \rangle} \bullet k_u \bullet B_{k_u+k_i,1} f^{(o)}(\right.$$

$$f^{(o)}(d+1) = \sum_{\vec{k}} \left[\frac{k_u P_{\vec{k}}}{\langle k_u \rangle} \bullet k_o B_{k_u+k_i,1} f^{(u)}(d) + \frac{k_i P_{\vec{k}}}{\langle k_i \rangle} \bullet k_o \bullet B_{k_u+k_i,1} f^{(o)}(d) \right]$$

- Gain ratio matrix:

$$\mathbf{R} = \sum_{\vec{k}} \begin{bmatrix} \frac{k_u P_{\vec{k}}}{\langle k_u \rangle} \bullet (k_u - 1) & \frac{k_i P_{\vec{k}}}{\langle k_i \rangle} \bullet k_u \\ \frac{k_u P_{\vec{k}}}{\langle k_u \rangle} \bullet k_o & \frac{k_i P_{\vec{k}}}{\langle k_i \rangle} \bullet k_o \end{bmatrix} \bullet B_{k_u+k_i,1}$$

- Spreading condition: max eigenvalue of $\mathbf{R} > 1$.

Global spreading condition:

- Useful change of notation for making results more general: write $P^{(u)}(\vec{k} \mid *) = \frac{k_u P_{\vec{k}}}{\langle k_u \rangle}$ and $P^{(i)}(\vec{k} \mid *) = \frac{k_i P_{\vec{k}}}{\langle k_i \rangle}$ where $*$ indicates the starting node's degree is irrelevant (no correlations).
- Also write $B_{k_u k_i,*}$ to indicate a more general infection probability, but one that does not depend on the edge's origin.
- Now have, for the example of mixed, uncorrelated random networks:

$$\mathbf{R} = \sum_{\vec{k}} \begin{bmatrix} P^{(u)}(\vec{k} \mid *) \bullet (k_u - 1) & P^{(i)}(\vec{k} \mid *) \bullet k_u \\ P^{(u)}(\vec{k} \mid *) \bullet k_o & P^{(i)}(\vec{k} \mid *) \bullet k_o \end{bmatrix} \bullet B_{k_u k_i,*}$$

Summary of contagion conditions for uncorrelated networks:

- I. Undirected, Uncorrelated—$f(d+1) = \mathbf{f}(d)$:

$$\mathbf{R} = \sum_{k_u} P^{(u)}(k_u \mid *) \bullet (k_u - 1) \bullet B_{k_u,*}$$

- II. Directed, Uncorrelated—$f(d+1) = \mathbf{f}(d)$:

$$\mathbf{R} = \sum_{k_i,k_o} P^{(i)}(k_i, k_o \mid *) \bullet k_o \bullet B_{k_i,*}$$

- III. Mixed Directed and Undirected, Uncorrelated—

$$\begin{bmatrix} f^{(u)}(d+1) \\ f^{(o)}(d+1) \end{bmatrix} = \mathbf{R} \begin{bmatrix} f^{(u)}(d) \\ f^{(o)}(d) \end{bmatrix}$$

$$\mathbf{R} = \sum_{\vec{k}} \begin{bmatrix} P^{(u)}(\vec{k} \mid *) \bullet (k_u - 1) & P^{(i)}(\vec{k} \mid *) \bullet k_u \\ P^{(u)}(\vec{k} \mid *) \bullet k_o & P^{(i)}(\vec{k} \mid *) \bullet k_o \end{bmatrix} \bullet B_{k_u k_i,*}$$

Correlated version:

- Now have to think of transfer of infection from edges emanating from degree $\vec{k}'$ nodes to edges emanating from degree $\vec{k}$ nodes.
- Replace $P^{(i)}(\vec{k} \mid *)$ with $P^{(i)}(\vec{k} \mid \vec{k}')$ and so on.
- Edge types are now more diverse beyond directed and undirected as originating node type matters.
- Sums are now over $\vec{k}'$.

Summary of contagion conditions for correlated networks:

- IV. Undirected, Correlated—$f_{k_u}(d+1) = \sum_{k_u'} R_{k_u k_u'} f_{k_u'}(d)$

$$R_{k_u k_u'} = P^{(u)}(k_u \mid k_u') \bullet (k_u - 1) \bullet B_{k_u k_u'}$$

- V. Directed, Correlated—$f_{k_i k_o}(d+1) = \sum_{k_i',k_o'} R_{k_i k_o k_i' k_o'} f_{k_i' k_o'}(d)$

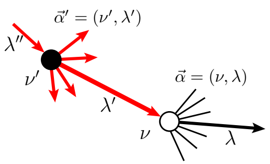
$$R_{k_i k_o k_i' k_o'} = P^{(i)}(k_i, k_o \mid k_i', k_o') \bullet k_o \bullet B_{k_i k_o k_i' k_o'}$$

- VI. Mixed Directed and Undirected, Correlated—

$$\begin{bmatrix} f^{(u)}_{\vec{k}}(d+1) \\ f^{(o)}_{\vec{k}}(d+1) \end{bmatrix} = \sum_{\vec{k}'} \mathbf{R}_{\vec{k}\vec{k}'} \begin{bmatrix} f^{(u)}_{\vec{k}'}(d) \\ f^{(o)}_{\vec{k}'}(d) \end{bmatrix}$$

$$\mathbf{R}_{\vec{k}\vec{k}'} = \begin{bmatrix} P^{(u)}(\vec{k} \mid \vec{k}') \bullet (k_u - 1) & P^{(i)}(\vec{k} \mid \vec{k}') \bullet k_u \\ P^{(u)}(\vec{k} \mid \vec{k}') \bullet k_o & P^{(i)}(\vec{k} \mid \vec{k}') \bullet k_o \end{bmatrix} \bullet B_{\vec{k}\vec{k}'}$$

Full generalization:



$$f_{\vec{\alpha}}(d+1) = \sum_{\vec{\alpha}'} R_{\vec{\alpha}\vec{\alpha}'} f_{\vec{\alpha}'}(d)$$

$R_{\vec{\alpha}\vec{\alpha}'}$ is the gain ratio matrix and has the form:

$$R_{\vec{\alpha}\vec{\alpha}'} = P_{\vec{\alpha}\vec{\alpha}'} \bullet k_{\vec{\alpha}\vec{\alpha}'} \bullet B_{\vec{\alpha}\vec{\alpha}'}.$$

- $P_{\vec{\alpha}\vec{\alpha}'}$ = conditional probability that a type λ' edge emanating from a type ν' node leads to a type ν node.
- $k_{\vec{\alpha}\vec{\alpha}'}$ = potential number of newly infected edges of type λ emanating from nodes of type ν.
- $B_{\vec{\alpha}\vec{\alpha}'}$ = probability that a type ν node is eventually infected by a single infected type λ' link arriving from a neighboring node of type ν'.
- Generalized contagion condition:

$$\max|\mu| : \mu \in \sigma(\mathbf{R}) > 1$$

- As we saw earlier, the triggering probability for simple contagion on random networks can be determined with a straightforward physical argument.
- Two good things:

$$Q_{\text{trig}} = \sum_{k=0}^{\infty} \frac{kP_k}{\langle k \rangle} \bullet B_{k1} \bullet \left[1 - \left(1 - Q_{\text{trig}}\right)^{k-1} \right],$$

$$P_{\text{trig}} = S_{\text{trig}} = \sum_k P_k \bullet \left[1 - (1 - Q_{\text{trig}})^k\right].$$

- Equivalent to result found via the eldritch route of generating functions.
- Generating functions arguably make some kinds of calculations easier (but perhaps we don't care about component sizes that much).
- On the other hand, a plainspoken physical argument helps us generalize to correlated networks more easily.

Summary of triggering probabilities for uncorrelated networks: [3]

- I. Undirected, Uncorrelated—

$$Q_{\text{trig}} = \sum_{k'_{\text{u}}} P^{(\text{u})}(k'_{\text{u}} \,|\, \cdot) B_{k'_{\text{u}}1} \left[1 - (1 - Q_{\text{trig}})^{k'_{\text{u}}-1}\right]$$

$$P_{\text{trig}} = S_{\text{trig}} = \sum_{k'_{\text{u}}} P(k'_{\text{u}}) \left[1 - (1 - Q_{\text{trig}})^{k'_{\text{u}}}\right]$$

- II. Directed, Uncorrelated—

$$Q_{\text{trig}} = \sum_{k'_{\text{i}}, k'_{\text{o}}} P^{(\text{u})}(k'_{\text{i}}, k'_{\text{o}} | \cdot) B_{k'_{\text{i}}1} \left[1 - (1 - Q_{\text{trig}})^{k'_{\text{o}}}\right]$$

$$S_{\text{trig}} = \sum_{k'_{\text{i}}, k'_{\text{o}}} P(k'_{\text{i}}, k'_{\text{o}}) \left[1 - (1 - Q_{\text{trig}})^{k'_{\text{o}}}\right]$$

Summary of triggering probabilities for uncorrelated networks:

- III. Mixed Directed and Undirected, Uncorrelated—

$$Q^{(\text{u})}_{\text{trig}} = \sum_{\vec{k}'} P^{(\text{u})}(\vec{k}' | \cdot) B_{\vec{k}'1} \left[1 - (1 - Q^{(\text{u})}_{\text{trig}})^{k'_{\text{u}}-1}(1 - Q^{(\text{o})}_{\text{trig}})^{k'_{\text{o}}}\right]$$

$$Q^{(\text{o})}_{\text{trig}} = \sum_{\vec{k}'} P^{(\text{i})}(\vec{k}' | \cdot) B_{\vec{k}'1} \left[1 - (1 - Q^{(\text{u})}_{\text{trig}})^{k'_{\text{u}}}(1 - Q^{(\text{o})}_{\text{trig}})^{k'_{\text{o}}}\right]$$

$$S_{\text{trig}} = \sum_{\vec{k}'} P(\vec{k}') \left[1 - (1 - Q^{(\text{u})}_{\text{trig}})^{k'_{\text{u}}}(1 - Q^{(\text{o})}_{\text{trig}})^{k'_{\text{o}}}\right]$$

Summary of triggering probabilities for correlated networks:

- IV. Undirected, Correlated— $Q_{\text{trig}}(k_{\text{u}}) = \sum_{k'_{\text{u}}} P^{(\text{u})}(k'_{\text{u}} \,|\, k_{\text{u}}) B_{k'_{\text{u}}1} \left[1 - (1 - Q_{\text{trig}}(k'_{\text{u}}))^{k'_{\text{u}}-1}\right]$

$$S_{\text{trig}} = \sum_{k'_{\text{u}}} P(k'_{\text{u}}) \left[1 - (1 - Q_{\text{trig}}(k'_{\text{u}}))^{k'_{\text{u}}}\right]$$

- V. Directed, Correlated— $Q_{\text{trig}}(k_{\text{i}}, k_{\text{o}}) = \sum_{k'_{\text{i}}, k'_{\text{o}}} P^{(\text{u})}(k'_{\text{i}}, k'_{\text{o}} | k_{\text{i}}, k_{\text{o}}) B_{k'_{\text{i}}1} \left[1 - (1 - Q_{\text{trig}}(k'_{\text{i}}, k'_{\text{o}}))^{k'_{\text{o}}}\right]$

$$S_{\text{trig}} = \sum_{k'_{\text{i}}, k'_{\text{o}}} P(k'_{\text{i}}, k'_{\text{o}}) \left[1 - (1 - Q_{\text{trig}}(k'_{\text{i}}, k'_{\text{o}}))^{k'_{\text{o}}}\right]$$

Summary of triggering probabilities for correlated networks:

- VI. Mixed Directed and Undirected, Correlated—

$$Q^{(\text{u})}_{\text{trig}}(\vec{k}) = \sum_{\vec{k}'} P^{(\text{u})}(\vec{k}' | \vec{k}) B_{\vec{k}'1} \left[1 - (1 - Q^{(\text{u})}_{\text{trig}}(\vec{k}'))^{k'_{\text{u}}-1}(1 - Q^{(\text{o})}_{\text{trig}}(\vec{k}'))^{k'_{\text{o}}}\right]$$

$$Q^{(\text{o})}_{\text{trig}}(\vec{k}) = \sum_{\vec{k}'} P^{(\text{i})}(\vec{k}' | \vec{k}) B_{\vec{k}'1} \left[1 - (1 - Q^{(\text{u})}_{\text{trig}}(\vec{k}'))^{k'_{\text{u}}}(1 - Q^{(\text{o})}_{\text{trig}}(\vec{k}'))^{k'_{\text{o}}}\right]$$

$$S_{\text{trig}} = \sum_{\vec{k}'} P(\vec{k}') \left[1 - (1 - Q^{(\text{u})}_{\text{trig}}(\vec{k}'))^{k'_{\text{u}}}(1 - Q^{(\text{o})}_{\text{trig}}(\vec{k}'))^{k'_{\text{o}}}\right]$$

Nutshell:

- Mixed, correlated random networks with undirected and directed edges form natural inclusive generalization of purely undirected and purely directed random networks.
- Spreading conditions and triggering probabilities of contagion processes can be determined using a direct, physical approach.
- These conditions can be generalized to arbitrary random networks with arbitrary node and edge types.
- More generalizations: bipartite affiliation graphs and multilayer networks.

References I

[1] M. Boguñá and M. Ángeles Serrano.
Generalized percolation in random directed networks.
Phys. Rev. E, 72:016106, 2005. pdf

[2] P. S. Dodds, K. D. Harris, and J. L. Payne.
Direct, phyiscally motivated derivation of the contagion condition for spreading processes on generalized random networks.
Phys. Rev. E, 83:056122, 2011. pdf

[3] K. D. Harris, J. L. Payne, and P. S. Dodds.
Direct, physically-motivated derivation of triggering probabilities for contagion processes acting on correlated random networks.
http://arxiv.org/abs/1108.5398, 2014.

References II

[4] M. E. J. Newman, S. H. Strogatz, and D. J. Watts.
Random graphs with arbitrary degree distributions and their applications.
Phys. Rev. E, 64:026118, 2001. pdf