Analysis of complex contagions in random multiplex networks

Osman Yağan and Virgil Gligor

Cascades in Complex Networks
- The past decade has witnessed a lot of research interest on dynamical processes in real-world complex networks.
- An interesting phenomenon that takes place in many such processes is the so-called information cascades.

Information Cascades: The spreading of an initially localized effect throughout the whole (or a very large part of) the network.

Examples:
- Diffusion of belief, norms, and innovations in social networks
- Disease contagion in human and animal populations
- Cascading failures in interdependent networks
- Global spread of computer viruses or worms on the Web

Our Focus: A class of dynamical processes: binary decisions with externalities.

Linear Threshold Model (Watts 2002)

Binary Decisions with Externalities
- Each individual must decide between two actions, e.g.,
  - To buy or not to buy a smartphone
  - To vote for Democrats or Republicans
- To join or not to join a dissident movement
- There is an incentive for individuals to coordinate their decisions with those of their immediate acquaintances.

Linear Threshold Model (Watts 2002)
- Nodes can be either one of the two states: active or inactive.
- Each node is initially given a threshold t drawn independently from Pθ(t).
- An inactive node with active neighbors and k ≥ t inactive neighbors will turn active if in fraction δ of active neighbors exceed t.
- Global Cascades: A linear fraction of nodes in the asymptotic limit eventually becomes active when an arbitrary node is made active. Wim (PNAS 99, 15764): Condition and Probability of global cascades (Gleeson & Cahalan [Phys. Rev. E 77, 46117]): Expected size is

Condition and Probability of Global Cascades

Simplex Networks
- Global cascade condition: Existence of a giant vulnerable component (GVC)
  - A node is vulnerable if its state can be changed by a single active neighbor, i.e., if i/δ < t.
  - Probability of global cascades: Fractional size of the extended giant vulnerable component
    - Extended GVC = Nodes that are connected to at least one node in GVC.

Multiplex Networks with Content-dependent Threshold Rule
- We need to define two notions of vulnerability:
  - A node is δ-vulnerable if it becomes active by a single active neighbor in W, i.e., if i/δ < t.
  - A node is δ-vulnerable if it becomes active by a single active neighbor in W, i.e., if i/δ < t.
- If G = A, the subgraph of vulnerable nodes forms a directed graph.
  - A potentially bi-directional F-link between nodes i and j will have the direction from i to j only if i > j (and vice versa for W-links).

Components of a directed network
- Out-component of a vertex is the set of vertices that are reachable from it.
- In-component of a vertex is the set of nodes that can reach that vertex.
- Giant out-component (GOUT): Set of nodes with infinite in-component.
- Giant in-component (GIN): Set of nodes with infinite out-component.
- Giant strongly-connected component (GSCC): Intersection of GIN and GOUT.

A subtle picture
- Condition for global cascades: Existence of GIN (Existence of nodes with infinite out-component)
  - Probability of global cascades: Fractional size of the extended giant in-component (EGIN)
    - Extended GIN = Nodes in GIN plus nodes that, once activated, can activate a node in GIN
  - Giant vulnerable component = GSCC
    - A set of vulnerable nodes s.t. activating any node (in this set) leads to the activation of all nodes in the set.

But, GIN may exist even if GSCC does not!
- A positive fraction of nodes have infinite out-component ⇒ GIN exists.
- The largest strongly connected component consists of two nodes ⇒ No GSCC.

Global cascades can take place even without a giant vulnerable component!
- Contradicts all previous models!

Network Model (An overlay social-physical network)
- Let r = 2, i.e., assume that there are only two link types.
- F: Random network of type-1 links with degree distribution ρf.
- W: Random network of type-2 links with degree distribution ρw.
- H: The overlay network F U W with a colored degree distribution ρh.
  - With ε := c/V an inactive node will become active with probability
    \[ p \left( \left( \frac{c}{V}, \frac{m}{k} \right)_k \right) \geq \varepsilon \Rightarrow \varphi(m, k) \]

Analytic Results

Branching Process for Exploring Out-Components
- Start by activating an arbitrary node, and then recursively reveal the largest number of vulnerable nodes reached and activated by exploring its neighbors.
  - \( q_i (\theta) \): Generating function for the integer number of nodes reached by following a type-i link, \( i = 1, 2 \)
  - \( p_{th} \): Probability that a node with color degree k is \( k \)-vulnerable, \( i = 1, 2 \)
  - \( G(\theta) \): Generating function for the finite number of nodes reached and activated.
    - Recursive relations:
      - Jacobian matrix
        \[ J(\theta) = \alpha \frac{\partial P_{th}^2}{\partial \theta} + \beta \frac{\partial P_{th}^2}{\partial \theta} - \gamma \frac{\partial P_{th}^2}{\partial \theta} \]
      - Global Cascade Condition: \( \varphi(\theta) > 1 \) (spectral radius)
        - Probability of Global Cascades, \( P_{out}^G = 1 - G(1) \)

Expected Cascade Size
- Construct a tree with a single node at the top level \( \ell = 0 \).
  - \( q_j (\theta) \): Probability that a node at level \( \ell \) is connected to its unique parent by a type-1 link, is active given that its parent is not. \( i = 1, 2 \)
    - Recursive relations:
      - \( q_j (\theta) = \frac{\partial P_{th}^2}{\partial \theta} \sum_{k=1}^{\infty} \sum_{k=1}^{\infty} F_{1} (\theta, k) \left( k - 1 \right) k \frac{\partial P_{th}}{\partial \theta} \]
  - \( \ell = 0, 1, 2 \)

Expected Size of Global Cascades: 5

Simulation Results (F, W ER with average degrees 21, 22)

- Excellent agreement between analysis and simulations!
- Content parameter \( c \) effects the range, probability and size of cascades!
- Parameter ranges that give positive values for \( S, P_{out} \) and \( S, \text{GSCC} \) are identical for \( c \neq 1 \).
- Global cascades without a giant vulnerable cluster is ruled out in our model, although this possibility exists in general.

For \( c \rightarrow \infty \)
- A positive fraction of nodes have infinite out-component ⇒ GIN exists.
- The largest strongly connected component consists of two nodes ⇒ No GSCC.

Threshold Rule with perceived proportion of active neighbors
- Consider a network where links can be of different types.
  - For a given content (e.g., rumor, product, political view), consider positive scalar \( c_1, c_2, \ldots, c_N \), such that \( \sum c_i = \tau \) quantifies the relative bias a type-i link has in spreading this particular content.
- Nodes switch state if their perceived proportion of active neighbors exceeds a threshold \( \tau \). Namely, an inactive node will become active if
  \[ c_1 m_1 + c_2 m_2 + \cdots + c_N m_N > \tau \]
  where \( m_i \) (resp. \( k_i \)) is the number of active neighbors (resp. number of neighbors) that the node is connected via a type-i link.

Content-dependent Threshold Model in Multiplex Networks

Our Motivation:
- Most existing works consider simple networks that have only a single type of link.
- But, individuals engage in different types of relationships; e.g., see the snapshot from Facebook on the right.
- Multiplex networks (multiple link types)
- Each link type may play a different role in different cascade processes
  - A virus game would be more likely to be promoted among high school classmates rather than among family members; the situation would be exactly the opposite in the case of a new cleaning product.
- Cascading failures in interdependent networks: Power links are vulnerable to natural hazards, computer links are vulnerable to viruses.