Network Science: Diffusion, Percolation, Tipping Points, Contagion and Cascades

Ozalp Babaoglu
Dipartimento di Informatica — Scienza e Ingegneria
Università di Bologna
www.cs.unibo.it/babaoglu/

Diffusion

- Viral spread of diseases, information, ideas — simple diffusion (contagion)
- Spread of new technologies, behaviors, opinions, fads, fashion — complex diffusion (peer-effects)
- Choices, decisions — games on networks (cascades)

Spatial networks

- Implicit networks that arise due to geographic proximity
- Nodes: individuals, edges: physical proximity
- Similar to Kleinberg’s small-world model: each node connected to its four compass neighbors

Simple diffusion in spatial networks

- Spread of forest fire in a two-dimensional grid
- Single “forestation” parameter sets the probability of a grid position being “forest” or “parking lot”
- Fire starts at a random grid position and spreads to all neighbors
- Observe the “spread” of diffusion (fire) as a function of forestation density
Simple diffusion in spatial networks

- If forestation probability of a grid position is $p$, then probability of it being a parking lot is $(1-p)$
- Fire will spread to all nodes in the connected component of the network containing the source node
- Run Library/Earth Science/Fire

Tipping phenomenon

- An abrupt change occurs in going from 57% to 62% forestation — the percentage of burnt forest suddenly increases from small values to almost 100%
- Sudden, massive increase in diffusion is a “tipping” phenomenon (also known as “threshold” or “critical” phenomena)
- Similar to the formation of a giant component in the ER model as the edge probability is increased

Tipping phenomenon

- “The last straw that broke the camel's back”
- Non-linear relation with a discontinuity
- What property of the camel exhibits a tipping phenomenon?

Not tipping phenomena

- Exponential growth

World Population

<table>
<thead>
<tr>
<th>Year</th>
<th>Population (billions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>3000</td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td></td>
</tr>
<tr>
<td>5000</td>
<td></td>
</tr>
</tbody>
</table>

Wikipedia Articles

<table>
<thead>
<tr>
<th>Year</th>
<th>Articles (hundred thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>0</td>
</tr>
<tr>
<td>2002</td>
<td>0</td>
</tr>
<tr>
<td>2003</td>
<td>0</td>
</tr>
<tr>
<td>2004</td>
<td>0</td>
</tr>
<tr>
<td>2005</td>
<td>0</td>
</tr>
<tr>
<td>2006</td>
<td>0</td>
</tr>
</tbody>
</table>
Can we get from top to bottom touching only grey squares? 
Analogy to water “percolating” through coffee grinds or oil seeping into the ground

Percolation depends on the “density” of the coffee grinds
Random model where each square is grey (empty) with probability $p$ and brown (coffee) with probability $1-p$

Probability of percolation

NetLogo Library/Earth Science/Fire and Percolation demos

Percolation

Simple Diffusion (Contagion) in networks
SI model

Diffusion or contagion can be formulated on any arbitrary network (not just grids)
Population divided into two groups
- Susceptible (S)
- Infected (I)
To make the model more realistic, add a parameter “infection rate” — determines the probability of disease spreading from an infected node to a susceptible node
- Susceptible-Infected (SI) model
- NetLogo ERDiffusion (random network)
- NetLogo BADiffusion (preferential attachment network)
Simple diffusion in networks

SIS model

- Susceptible-Infected-Susceptible (SIS) model
- After being infected, individuals remain susceptible
- Appropriate for modeling recurring diseases
- NetLogo “SmallWorldDiffusionSIS” (small-world network)

SIR model

- Susceptible-Infected-Resistant (SIR) model
- Allows resistance or immunity to be gained after infection
- Add a new “Recovered” state and in addition to the infection rate, add a new parameter “recovery rate”
- We can also add a parameter “gain resistance” if infection does not guarantee resistance with certainty
- NetLogo “Library/Networks/Virus on a Network”

Two-parameter SIR model:

- $a$: infection rate
- $b$: recovery rate

Population divided into three groups:

- Let $S(t)$ be the number of susceptible at time $t$
- Let $I(t)$ be the number of infected at time $t$
- Let $R(t)$ be the number of recovered at time $t$

\[
\frac{dS}{dt} = -aSI \\
\frac{dI}{dt} = aSI - bI \\
\frac{dR}{dt} = bI
\]
Look at \( \frac{dl}{dt} = aSI - bl \) and compare it to zero.

- \( aSI - bl < 0 \) or \( aS - b < 0 \) or \( \frac{aS}{b} < 1 \)

\( \frac{aS}{b} \) is known as \( R_0 \) and is a critical parameter in epidemiology.

We can decrease \( R_0 \) by
- decreasing the infection rate \( a \) (washing hands, social distancing),
- by decreasing the susceptible population \( S \) (vaccination),
- by increasing the recovery rate \( b \) (better health care, usually difficult)

Around \( t=0 \) assume \( S(t) \) is constant.

Solve the differential equation \( \frac{dl}{dt} = l(aS_0 - b) \)

\( I(t) = e^{(aS_0 - b)t} \)

In other words, if \( aS_0 - b > 0 \) (or \( R_0 > 1 \)) the infected population will start to grow exponentially and we have an epidemic.

If \( R_0 < 1 \) the infected population will decay and there will not be an epidemic.
Simple diffusion in networks

SIR model

- Covid-19 confirmed cases as of 26 March 2020

Country by country: how coronavirus case trajectories compare
Cumulative number of confirmed cases, by number of days since 100th case

SEIR model

- Susceptible → Exposed → Infected → Removed

Peer-effects

- The "infection rate" is often not a constant but depends on how many nodes are already infected
- Captures the notion of "peer-effects" — we are more likely to adopt the behavior or choices of our peers
- The greater the number of my peers who dress, talk, walk, eat or vote one way, the higher the probability that I will dress, talk, walk, eat or vote that way too

Simple diffusion in networks

Epidemic Calculator
http://gabgoh.github.io/COVID/index.html
Peer-effects
S-shape curves

Adoption rate of hybrid corn in three states

- Initial adopters
- Accelerated adoption
- Saturation

Griliches (1957): Hybrid Corn Diffusion S-Shape, Spatial Pattern...

- Kentucky
- Wisconsin
- Iowa

Adoption rate of new drugs by doctors

- Named by 0 others
- Named by 1, 2 others
- Named by 3+ others

Peer-effects
Bass model

- No explicit network
- Two states/behaviors: 0 and 1
- States are irreversible — cannot switch back and forth
- Let $F(t)$ denote the fraction of population who have adopted state 1 at time $t$
- Let $p$ denote the rate of spontaneous adoption
- Let $q$ denote the rate of peer-effect adoption
- How does the fraction of adoption vary with time?

$$\frac{dF(t)}{dt} = (p + qF(t))(1 - F(t))$$

- When $F(t)$ nears 1, $\frac{dF(t)}{dt}$ nears 0
- When $F(t)=0$, $\frac{dF(t)}{dt}=p$
- When $F(t)=\epsilon$ for some $\epsilon$, $\frac{dF(t)}{dt}=(p+q\epsilon)(1-\epsilon)$
- To get initial convexity, $(p+q\epsilon)(1-\epsilon)$ must be greater than $p$
- Or $q(1-\epsilon)>p$ and for small $\epsilon$, this is equivalent to $q > p$
- Thus, we get the “S-shape” if $q > p$
Peer-effects Bass model

where

\[ F(t) = \frac{(1 - e^{-(p+q)t})}{(1 + \frac{q}{p}e^{-rt})} \]

Complex diffusion in networks

- **Simple diffusion**: just one node sufficient to “infect” another node (with some fixed probability)
- **Peer-effects diffusion**: probability of infection depends on number of nodes already infected
- **Complex diffusion**: a node being “infected” depends on choices or decisions made by the node
- Appropriate for modeling adoption of new technologies, behaviors, opinions, fashion trends, etc.

Example: Suppose a node will adopt a behavior (mohawk haircut) only if two of its neighbors have adopted the behavior
- NetLogo “SmallWorldDiffusionComplex” (adopt if two neighbors share opinion)
- In general, the decision to adopt or not could be much more complex
- Can be modeled as a “network coordination game”

Complex diffusion in networks

Network coordination game

- Two possible behaviors: \( A \) and \( B \)
- Two nodes that are neighbors have an incentive to adopt the same behavior
  - Units of measure (Metric vs Imperial)
  - Sports (Basketball vs Soccer)
  - Social networking (Facebook vs Google+)
- Payoff matrix

<table>
<thead>
<tr>
<th></th>
<th>( A )</th>
<th>( B )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>( a, a )</td>
<td>( 0, 0 )</td>
</tr>
<tr>
<td>( B )</td>
<td>( 0, 0 )</td>
<td>( b, b )</td>
</tr>
</tbody>
</table>
Network coordination game

- Each node plays a copy of the game with each of its neighbors
- The utility of a node is the sum of the payoffs obtained from the individual games
- Consider a node $v$ with $d$ neighbors of which a fraction $p$ have chosen $A$

![Diagram of network coordination game]

- If node $v$ chooses $A$, its utility will be $p da$
- If it chooses $B$, its utility will be $(1-p) db$
- Thus, $A$ is a better choice if $p da \geq (1-p) db$ or $p \geq b/(a+b)$
- In other words, if at least $b/(a+b)$ fraction of $v$’s neighbors chose $A$, then $v$ should choose $A$ as well
- If $b/(a+b)$ is small, $A$ is the more attractive choice
- If $b/(a+b)$ is large, $B$ is the more attractive choice

Cascading behavior

- There are two obvious equilibria: everyone chooses $A$, or everyone chooses $B$
- Suppose that the network is in the second equilibrium: everyone has chosen $B$
- Can the network be “tipped over” to the other equilibrium by flipping the choices of a small number of “initial adopter” nodes?
- Answer depends on the network structure, the ratio $b/(a+b)$ and the choice of initial adopters
- Initial adopters switch for reasons external to the game
- The other nodes continue to play the coordination game

Chain reaction of switches from decision $B$ to decision $A$ is called a cascade
- Cascades can be either *complete* — the entire network eventually switches to the other decision — or they can be *partial*
**Cascading behavior**

Application to viral marketing

- Partial cascades result in situations in which two (or more) decisions coexist
- Suppose $A$ and $B$ are two competing products or technologies and the manufacturer of $A$ wants to dominate the market (obtain a complete cascade)
- Manufacturer of $A$ has two possible strategies:
  - Make its product more "competitive" (increase $a$)
  - Pick very carefully the set of its initial adopters
  - Usually, modifying the network structure is not an option

Pursuing the first strategy, suppose the manufacturer of $A$ is able to increase $a$ from 3 to 4 (while $b$ remains at 2)
- The threshold for adopting $A$ reduces from $2/5$ to $1/3$
- In the example, the adoption of $A$ becomes a complete cascade

Pursuing the second strategy, suppose the manufacturer of $A$ is able to convince two additional nodes

- Convince two additional nodes
Cascading behavior
Application to viral marketing

- Or, pick a different set of initial adopters

\[ a=3 \]
\[ b=2 \]

Cascading behavior
Characterizing partial cascades

- Which feature of a network causes a cascade to stall rather than complete?

- **Definition:** A set of nodes is a cluster of density \( p \) if each node in the set has at least a \( p \) fraction of its network neighbors in the set.

Cascading behavior
Characterizing partial cascades

- Consider a network with threshold \( b/(a+b) \) for adopting decision \( A \) and a set of initial adopters of decision \( A \)

  (i) If the remaining network contains a cluster of density greater than \( 1-b/(a+b) \), then the set of initial adopters will not cause a complete cascade

  (ii) Moreover, whenever a set of initial adopters does not cause a complete cascade with threshold \( b/(a+b) \), the remaining network must contain a cluster of density greater than \( 1-b/(a+b) \)