Contagion in Networks

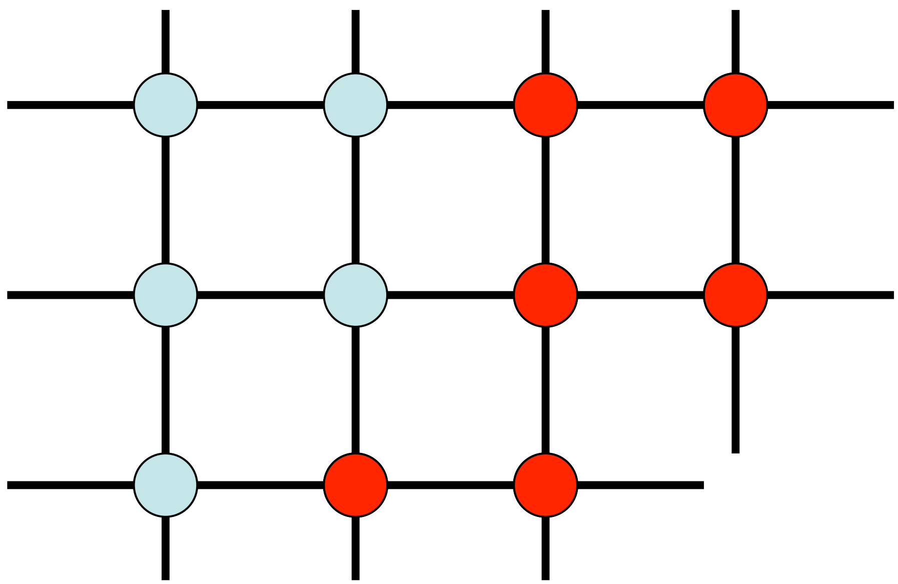
Networked Life
NETS 112
Fall 2015
Prof. Michael Kearns

Two Models of Network Formation

- Start with a grid, remove random fraction of vertices
  - “local” or “geographic” connectivity
- Start with N isolated vertices, add random edges
  - “long distance” connectivity
- Examine a deterministic contagion model
- Widespread infection occurs at “tipping point” of connectivity

“Mathematizing” the Forest Fire
(see Coursera “Contagion” video)

- Start with a regular 2-dimensional grid network
  - this represents a complete forest
- Delete each vertex (and all 4 of its edges) with probability 1-p
  - p is fraction of forest, 1-p is fraction of parking lots or clear-cut
- Choose a random remaining vertex v
  - this is my campsite
- Q: What is the expected size of v’s *connected component?*
  - i.e. the number of vertices reachable from v
  - this is how much of the forest is going to burn
- Observe a “tipping point” around p = 0.6



“Mathematizing” the Average Degree Demo
(see Coursera “Contagion” video)

- Let d be the desired average degree in a network of N vertices
- Then the total number of edges should be

$$e = d \times N / 2$$

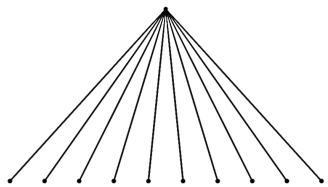
- Just start connecting random pairs of vertices until you have e edges
- Pick a random vertex v to infect
- What is the size of v’s connected component?
- Observe a “tipping point” around d=3

Some Remarks on the Demos

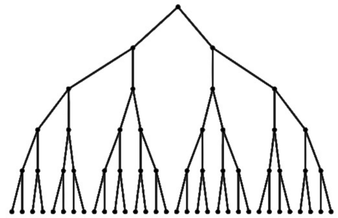
- Connectivity patterns were either *local* or *random*
  - will eventually formalize such models
  - what about other/more realistic structure?
- Tipping was inherently a *statistical* phenomenon
  - probabilistic nature of connectivity patterns
  - probabilistic nature of disease spread
  - model *likely* properties of a large *set* of possible outcomes
  - can model either inherent randomness or variability
- Formalizing tipping in the forest fire demo:
  - might let grid size N → infinity, look at fixed values of p
  - is there a threshold value q:
    - $p < q$ → expected fraction burned < 1/10
    - $p > q$ → expected fraction burned > 9/10

"Structural Virality"
Goel, Anderson, Hofman, Watts

- Every video, news story, image, or petition posted to Twitter over 12 months (1.4 B observations)
  - Restrict to "popular" cascades (> 100 RTs; ~350K events)
- For each event, can quantity its "structural virality"
  - Average Pairwise Shortest Path Length
  - Ranges from
    - ≈2 ("broadcast")
    - ~log(N) ("viral")
- For these "popular" events can ask:
  - What diversity do we see with respect to structure?
  - What is the relationship between size and structural virality?

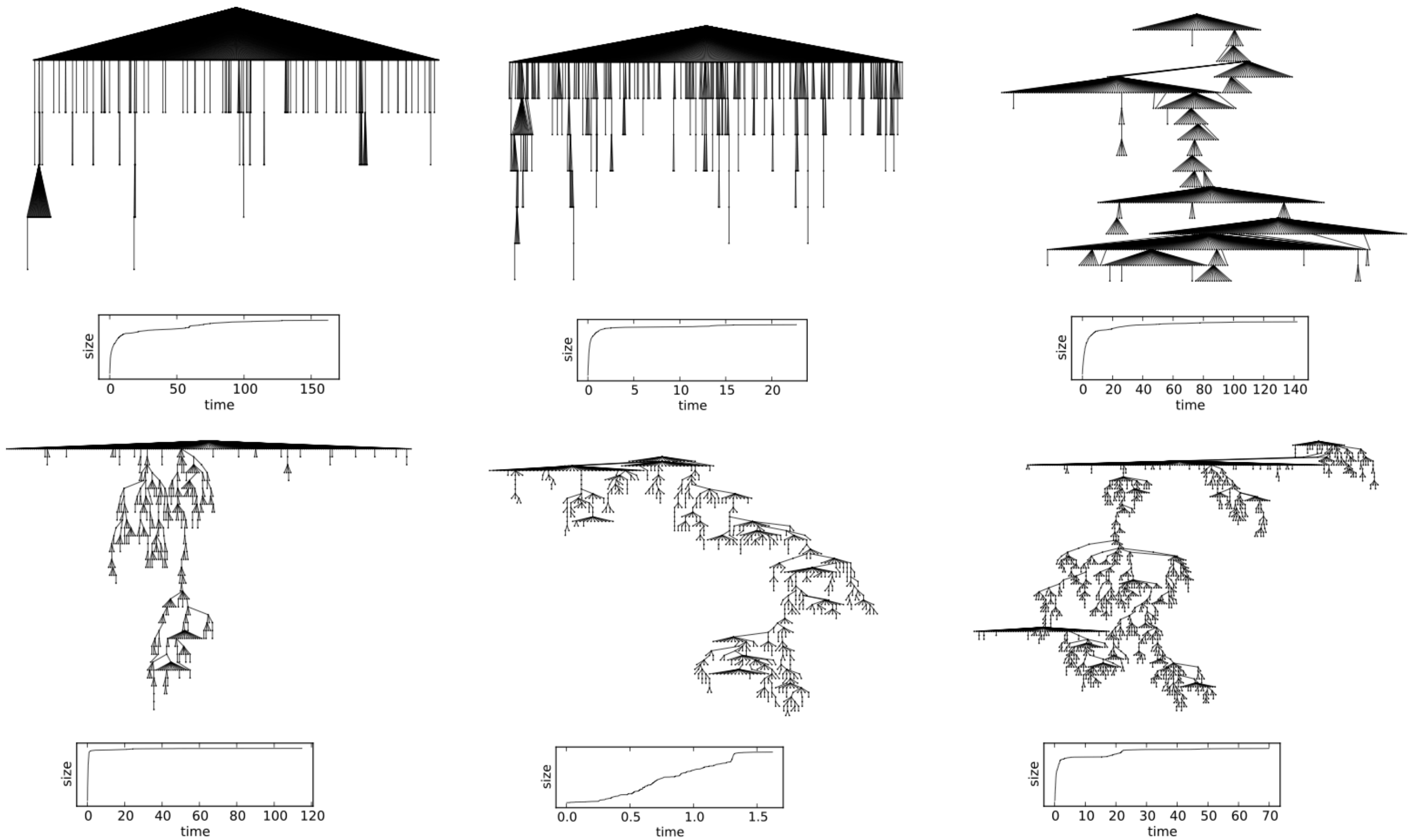


"Broadcast"

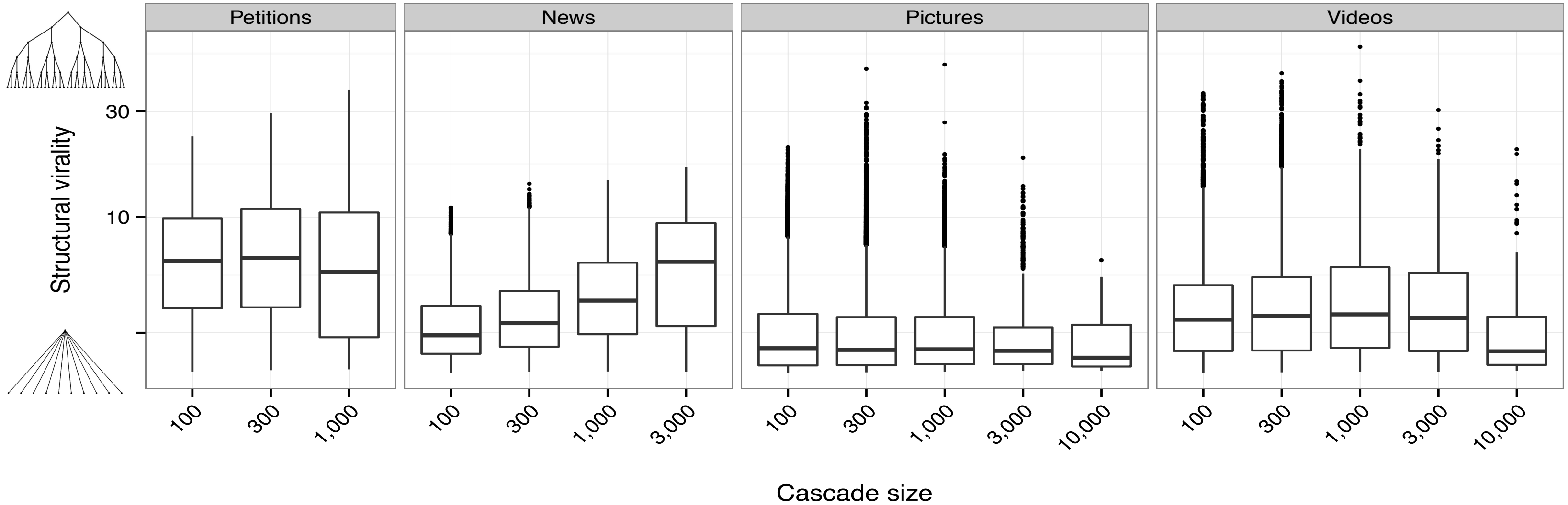


"Viral"

Diversity of Structural Virality

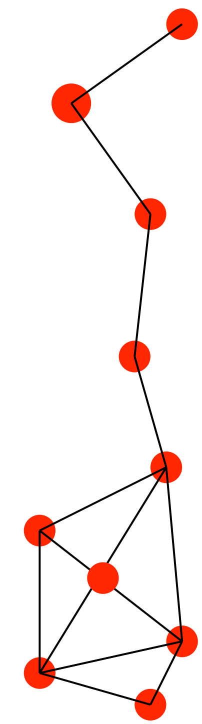


Popular ≠ Viral



Popularity driven mostly by the size of the largest broadcast

Structure and Dynamics Case Study: A "Contagion" Model of Economic Exchange



- Imagine an undirected, connected network of individuals
  - no model of network formation
- Start each individual off with some amount of currency
- At each time step:
  - each vertex divides their current cash equally among their neighbors
  - (or chooses a random neighbor to give it all to)
  - each vertex thus also *receives* some cash *from* its neighbors
  - repeat
- A *transmission* model of economic exchange --- no "rationality"
- Q: How does network structure influence outcome?
- A: As time goes to infinity:
  - vertex i will have fraction deg(i)/D of the wealth; D = sum of deg(i)
  - degree distribution *entirely* determines outcome!
  - "connectors" are the wealthiest
  - not obvious: consider two degree = 2 vertices…
- How does this outcome change when we consider more "realistic" dynamics?
  - e.g. we each have goods available for trade/sale, preferred goods, etc.
- What other processes have similar dynamics?
  - looking ahead: models for web surfing behavior