Network Analysis:

The Hidden Structures behind the Webs We Weave

17-213 / 17-668

Social Dynamics on Networks: Diffusion and Contagion

Tuesday, November 19, 2024

Patrick Park & Bogdan Vasilescu

Carnegie Mellon University
School of Computer Science

S3D
Software and Societal Systems Department

2-min Quiz, on Canvas

Dynamics on Social Networks

Things Spread through Networks

Information: News, ideas, knowledge

Preferences: predilections, cultural taste

Physiological / psychological states: Emotions, obesity, yawning

Socio-cultural artifacts: Customs, values, beliefs, norms, law, institutions

Macro-Structural Questions:

How can we quantitatively describe these spreading dynamics?

Can we predict the speed and magnitude of the spreading?

What explains these spreading dynamics?

Information Diffusion

Spread the word: Viral marketing

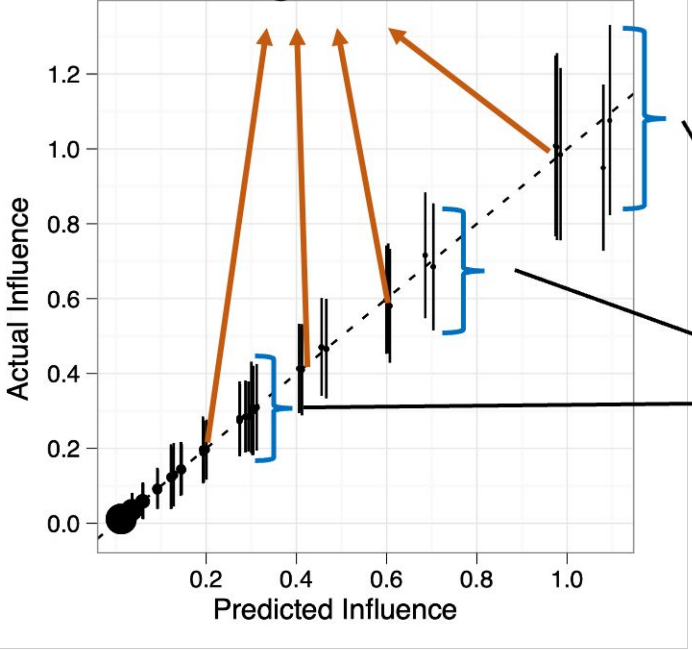
Question: Who should you target in a network to “maximize” information cascades?

- 74M separate diffusion events (Twitter retweets of URLs)
- Influence of the seed node: # of nodes in the diffusion tree
- Seed node’s attributes (followers, friends, tweets) and previous success of the seed node most predictive of average influence scores of the leaf nodes (clusters) in the regression tree

Answer: Hard to predict

Diffusion is difficult to predict

Highly accurate predictions of within-leaf average influence scores



-Regression tree model not so predictive of individual influence scores
-Weak effect of the nature of the content
-With these “null” results, the paper pivots to asking a slightly different question: Who should you target to “optimize” information cascades (i.e., introduce cost constraint)?

Variance within leaf (cluster) too large to predict individual influence scores

Source: Goel et al. 2015

Structural virality of diffusion

How do information cascades look like?

Structural virality (Wiener index)

- Average path length in a diffusion tree

$$\nu(T) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}$$

Recall

In a tree

Structural virality of diffusion

Structural virality (Wiener index)

- Average path length in a diffusion tree

$$\nu(T) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}$$

Structural virality of diffusion

Structural virality (Wiener index)

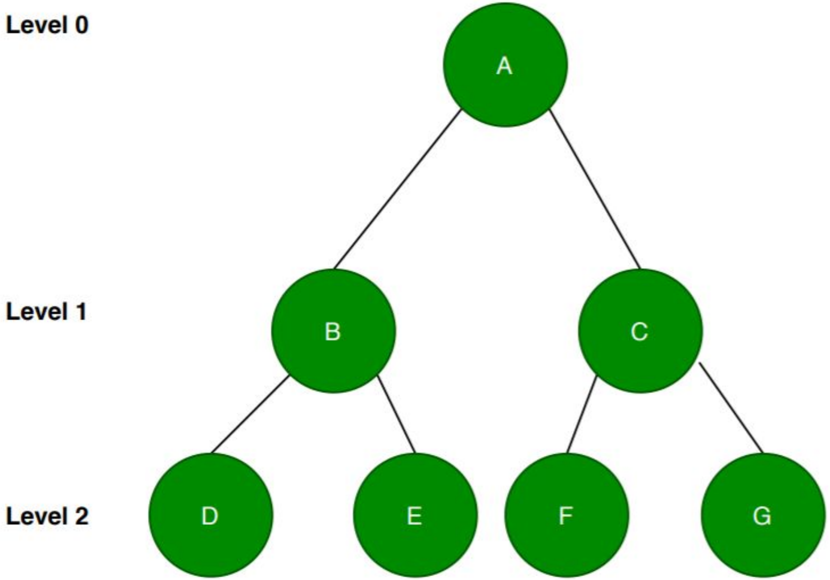
- Average path length in a diffusion tree

$$\nu(T) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}$$

Recall, *d* ~ Ln(*N*) / Ln<*k*>

In a complete binary tree

- N=2^0+2^1+...2^h
- Ln(N) ~ h * Ln(2)
- h ~ Ln(N) / Ln(2) → <k> = 2
- h ~ Ln(N) / Ln<k>
- *d* ~ *h*



Structural virality of diffusion

Examples of information cascade trees in increasing order of virality

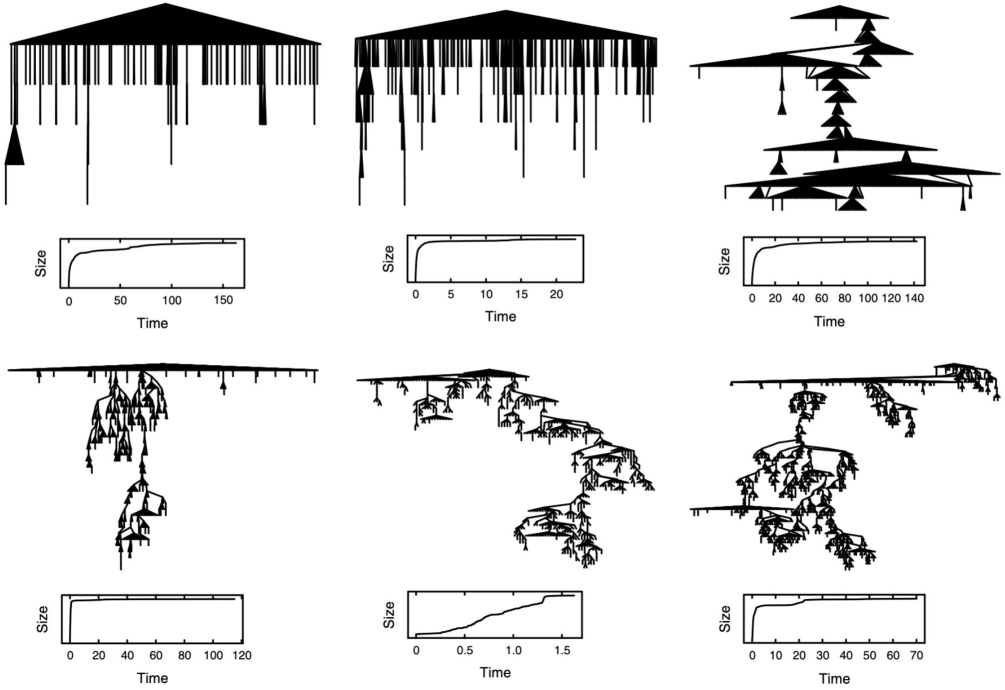
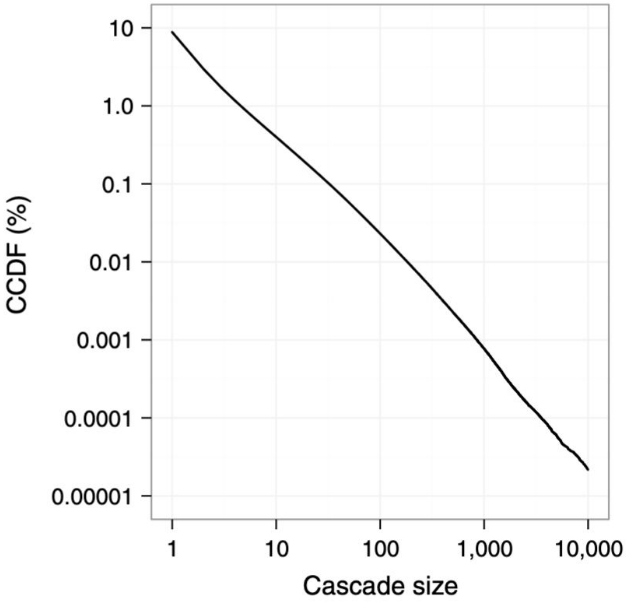


Figure 2 **Distribution of Cascade Sizes on a Log–Log Scale, Aggregated Across the Four Domains We Study: Videos, News, Pictures, and Petitions**

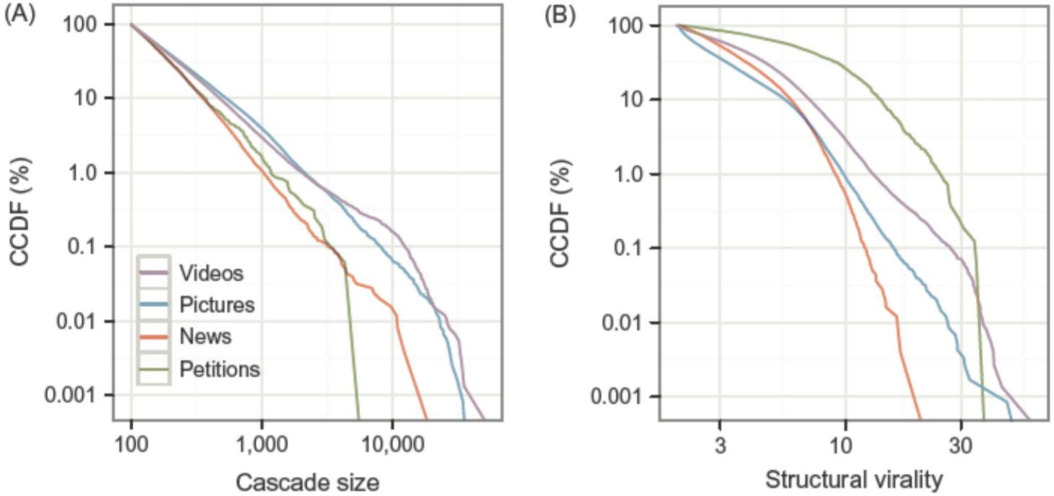


Note. CCDF, complementary cumulative distribution function.

Structural virality of diffusion

Does structural virality correlate with cascade size?

Figure 4 **Size and Structural Virality Distributions on a Log–Log Scale for Cascades Containing at Least 100 Adopters, Separated by Domain**



Note. CCDF, complementary cumulative distribution function.

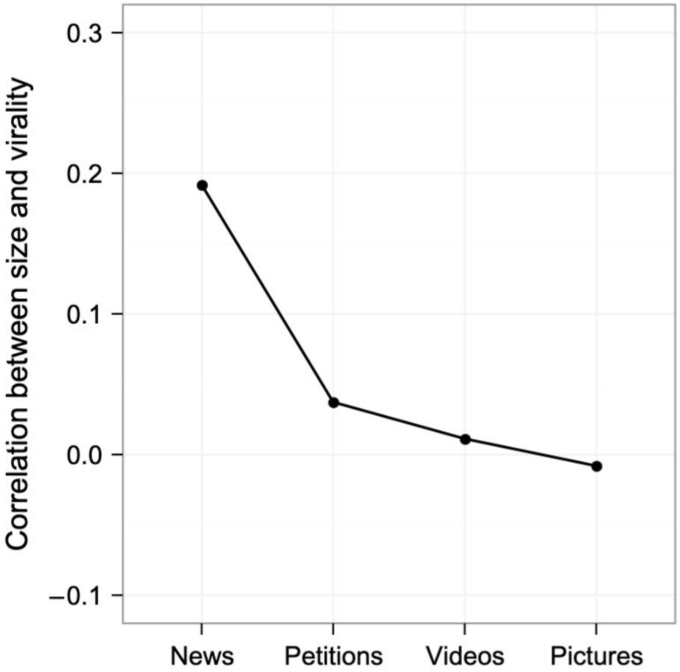
Structural virality of diffusion

Does structural virality correlate with cascade size?

- Not really

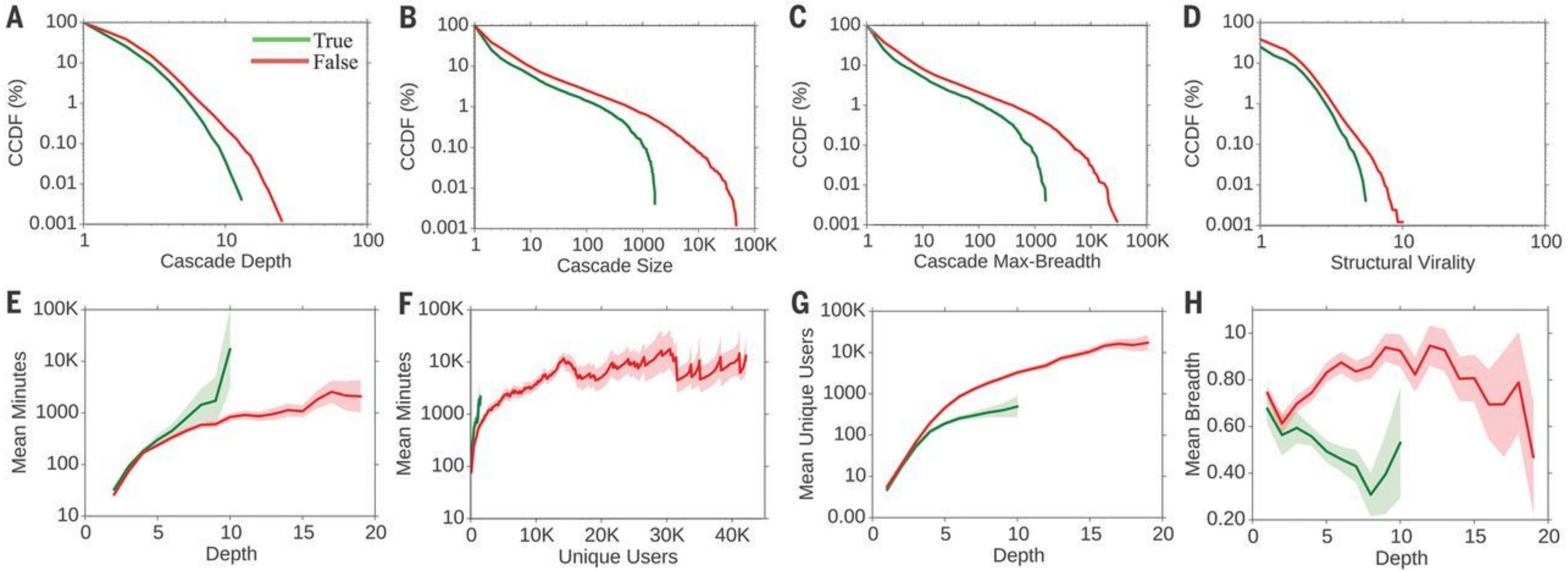
Predicting mass information diffusion is hard

Figure 6 Correlation Between Cascade Size (Popularity) and Structural Virality Across Four Domains



True vs. False information diffusion

False news diffuses much faster, reaches broader audience, and penetrates more deeply



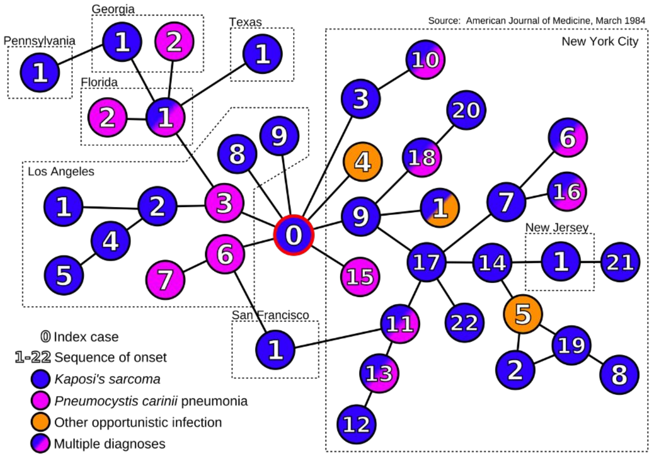
Complex Contagion: Costly spreading

Simple Contagion

A single contact leads to adoption/contagion (e.g., virus)

Spreads quickly in networks with low CPL (e.g., small-world)

Individual with a diverse egonetwork can "infect" disproportionately (e.g., super spreaders)



Dynamics of Behavioral Change

Model the effect of network structure on the spread and adoption of behaviors through network ties

Three Mechanisms of social adoption
-Common environmental influence

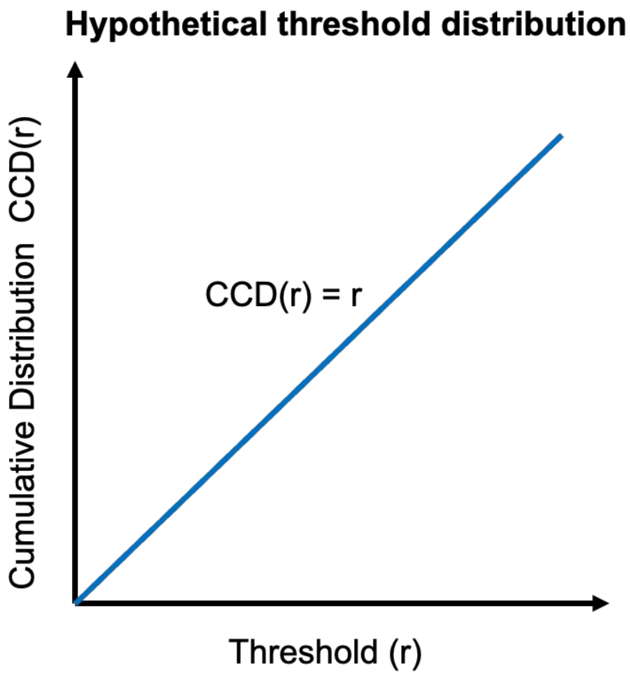
-Homophily (e.g., similar taste)

-Social influence

Very difficult to disentangle these mechanisms with observational data

(e.g., Framingham study of the spread of obesity)

Threshold models of adoption



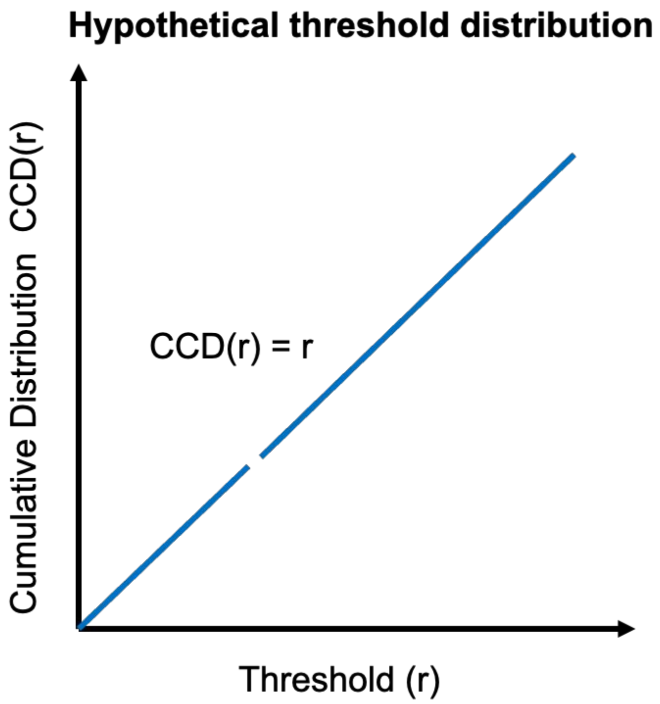
Some social behaviors require more than single exposure for adoption

- Individuals can have different levels of reluctance/resistance (thresholds)

- Variance in norms, preferences, utility lead to a distribution of thresholds

- Toy example: If an initial adoption occurs, adoption will reach 100% (saturation)

Threshold models of adoption



Sensitivity of collective behavior

- A negligible change to the threshold distribution can lead to vastly different equilibria

Threshold models of adoption

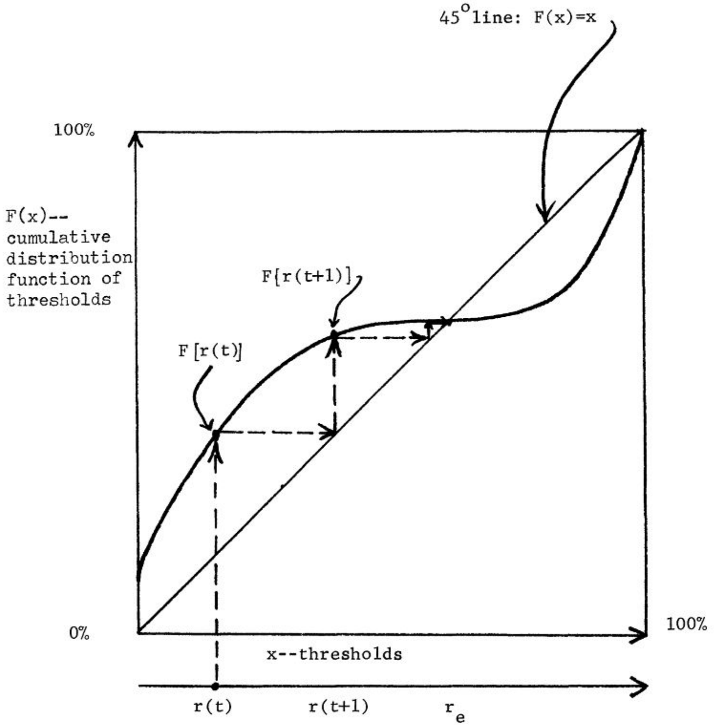


FIG. 1.—Graphical method of finding the equilibrium point of a threshold distribution. $r(t)$ = proportion having rioted by time t.

Some social behaviors require more than single exposure for adoption

- Assumption 1: People have perfect information about adoption at time *t*

- Assumption 2: Individual's threshold pertains to population adoption, not local adoption

Dynamics of Behavioral Change

Identification strategy: **experimental** approach

- Create two separate worlds, with and without social influence
- Observe adoption behavior in the two worlds
- Example: The Music Lab experiment

The Music Lab Experiment

Weak influence condition

Strong influence condition

Complex Contagion

Social contagion is an endogenous process:

- Homophily → adoption
- Embeddedness → adoption
- Tie strength → adoption

Similar people form strong ties

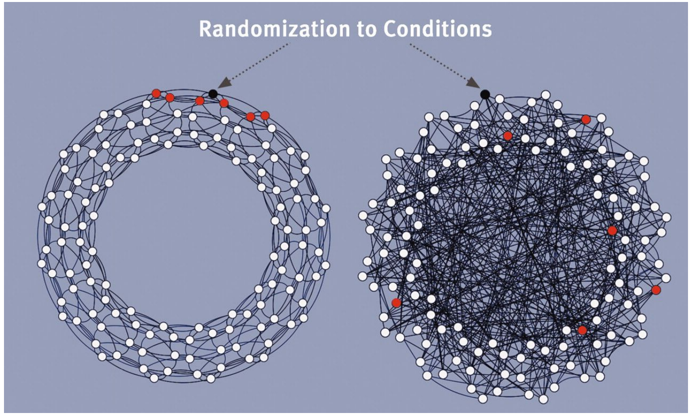
Embedded relations tend to be strong ties

Tie strength can potentially increase similarity

Tie strength can generate embedded relations

Result: **Difficult to estimate causal effect on adoption**

Complex Contagion: Randomized Experiment



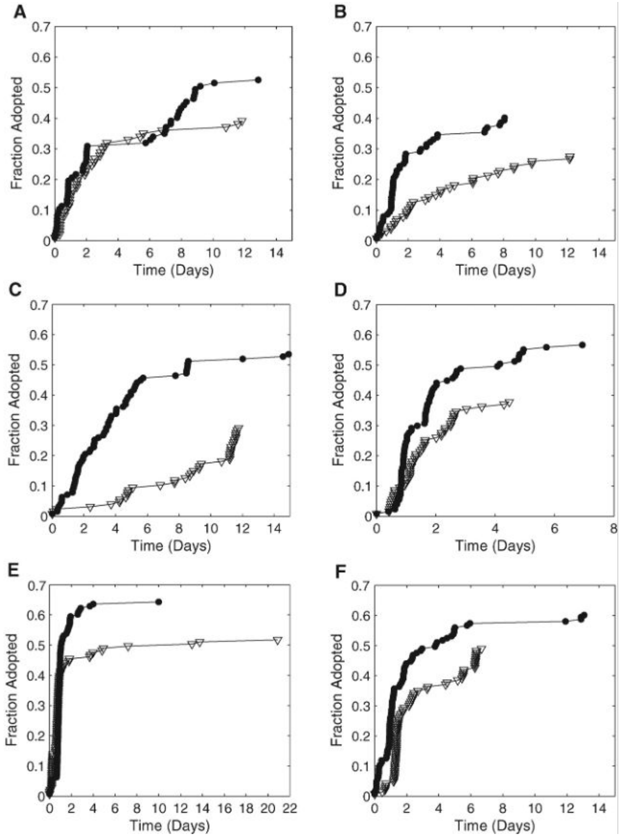
Adoption/infection probability increases with the number of neighbors who already adopted

Builds on the ideas of thresholds and social reinforcement

Initially studied as a simulation model (Centola and Macy 2007)

Centola reproduced the results through real-world experiments

Complex Contagion: Randomized Experiment



Adoption/infection probability increases with the number of neighbors who already adopted

Builds on the ideas of thresholds and social reinforcement

Initially studied as a simulation model (Centola and Macy 2007)

Centola reproduced the results through real-world experiments

Complex Contagion

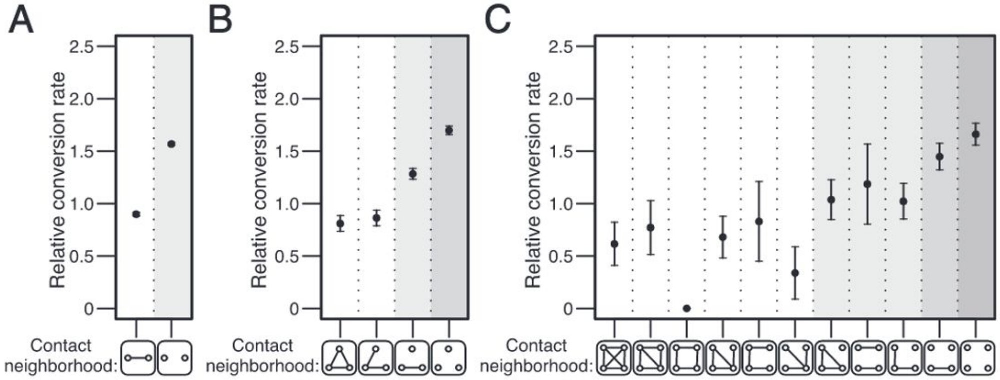


Fig. S2. Recruitment conversion for demographically homogeneous neighborhoods, as a function of (*A*) two-node, (*B*) three-node, and (*C*) four-node contact neighborhood graphs. The conversion scale is the same as for Fig. 1 in the main text. Error bars represent 95% confidence intervals.

Ugander et al. 2012

Open questions:

For a focal individual, is a closed or open triad more conducive to social contagion? (e.g., Facebook adoption study)

Opinion Dynamics on Networks: Why Liberals Drink Lattes

The Problem of Lifestyle Politics

Latte-drinking liberals and bird-hunting conservatives

Latte-liberal stereotype has a long history

Attribute-Based Explanations

Political ideology is correlated with lifestyle items in the General Social Survey

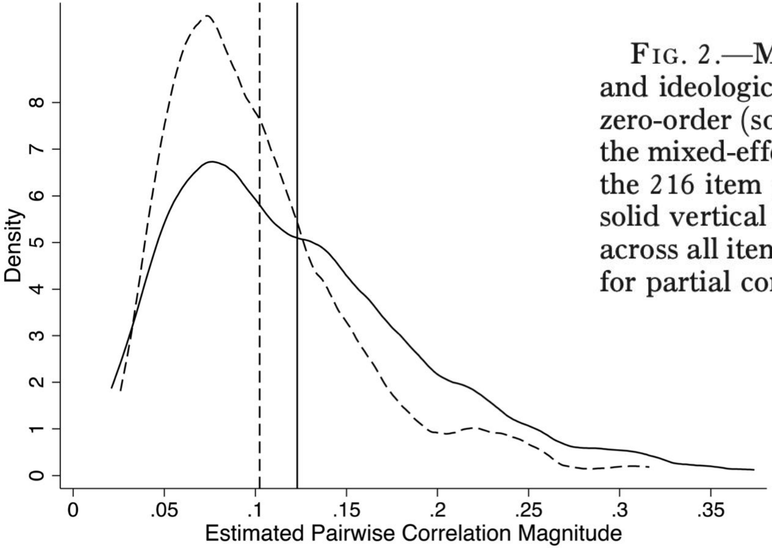


FIG. 2.—Magnitude of zero-order and partial correlation between GSS lifestyle items and ideological identity. Graphs plot Epanechnikov kernel density functions for both zero-order (solid lines) and partial (dashed lines) correlation magnitudes estimated from the mixed-effects model (see table A2 in the appendix). One value is plotted for each of the 216 item pairs. Time is set to 2010 to facilitate comparison across item pairs. The solid vertical reference line gives the mean predicted zero-order correlation magnitude across all item pairs in 2010 and the dashed vertical reference line gives the same value for partial correlation magnitude.

DellaPosta et al. 2015

The Problem of Lifestyle Politics

Latte-drinking liberals and bird-hunting conservatives

Lattes and bird-hunting have no inherent relationship with political orientation

Other examples: musical taste and political orientation

- Liberals are omnivorous: positive correlation with blues, reggae, jazz, rock
- Conservatives with stronger belief in religion vs. science

Q: How did we come to form these stereotypes?

Attribute-Based Explanations

Q: How did we come to form these stereotypes?

Attribute-based explanations:

- **Education**: People develop taste for certain lifestyles (e.g., classical music)
- **Economic status**: Certain lifestyles are costly
- **Occupation**: work that is complex, low supervision, and creative make people less conforming and liberal
- **Moral values**: care, fairness, liberty vs. loyalty, authority, sanctity
- **Psychological traits**: openness to new experience and cognitive complexity vs. need for certainty
- **Physiological differences**: Age, gender

Network Autocorrelation

Problem of attribute-based explanations:

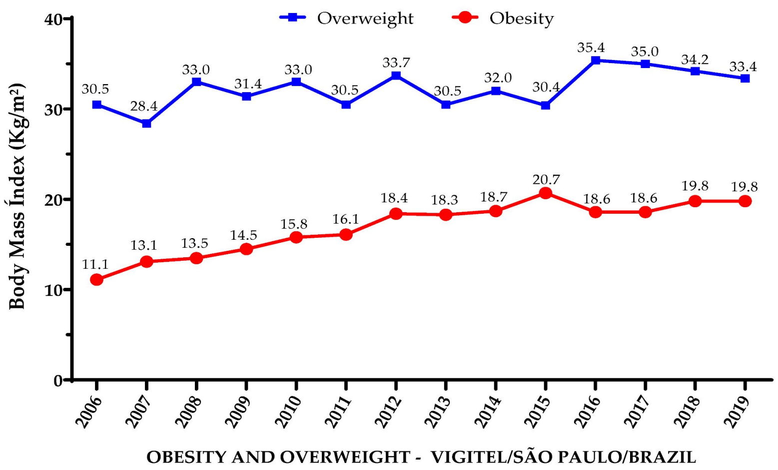
- Attribute-based explanations implicitly assume that individuals are social atoms
- Regression analysis of survey data assumes independent observations (individuals)

Before constructing elaborate explanations about lifestyle and politics, one must rule out the simplest explanation first: network autocorrelation

Network Autocorrelation

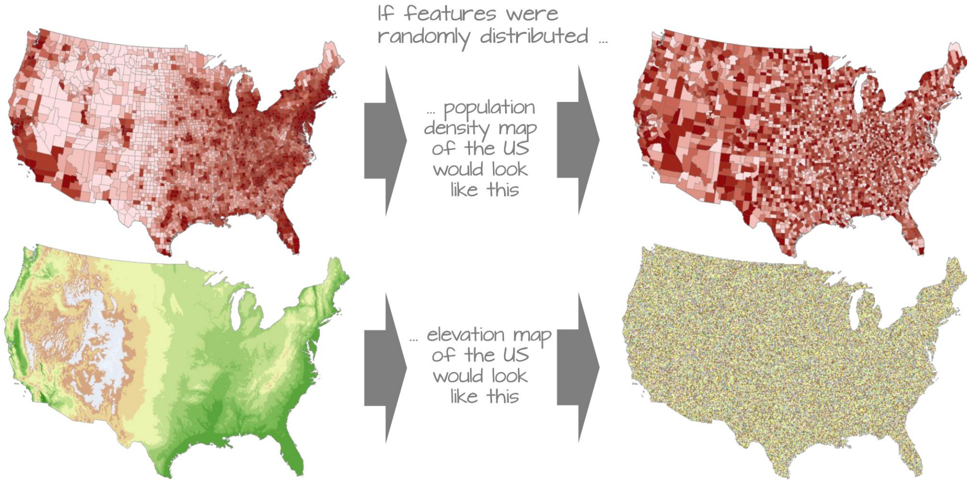
Autocorrelation: An observation is dependent on other observations, where this dependence increases with proximity in temporal, spatial, and network location.

Temporal Autocorrelation



Source: De Lima et al. 2024

Spatial Autocorrelation



Source: Manual Gimond Github

Network Autocorrelation

Network autocorrelation:

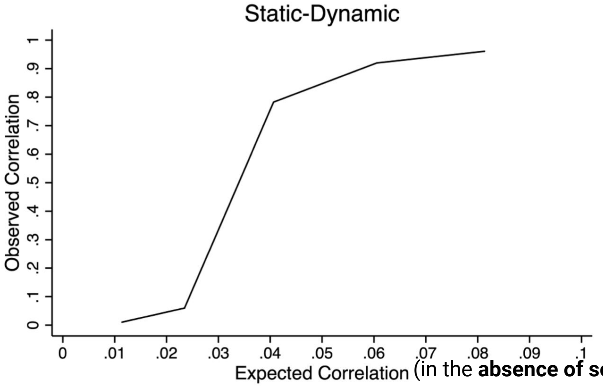
- People are influenced by network neighbors (e.g., peer approval)
- Herding effect when environmental uncertainty is high (i.e., follow the crowd)

Self-reinforcing dynamic of homophily and social influence explains lifestyle - politics correlation

- Similarity strengthens a social tie (homophily)
- The strengthened social tie leads to even higher similarity (social influence)
- Initially small correlations (stochastic noise) in politics and lifestyle preferences get amplified

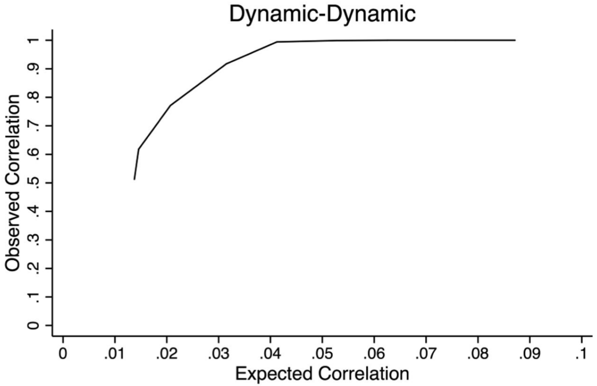
Network Autocorrelation

From simulations **with social influence**



(in the **absence of social influence**)

Static trait (e.g., gender, race) and dynamic trait (e.g., political belief)



Dynamic trait (e.g., education) and dynamic trait (e.g., political belief)

DellaPosta et al. 2015

Lifestyle Politics Are Correlations, Not Causations

Tim Walz: A bird-hunting Democrat

Summary

-