Contagion

Santa Fe Institute Summer School, 2009

Prof. Peter Dodds

Department of Mathematics & Statistics
Center for Complex Systems
Vermont Advanced Computing Center
University of Vermont

Outline

Introduction
Simple Disease Spreading Models
Background
Prediction
Social Contagion Models
Granovetter's model
Network version
Groups
Summary
Winning: it's not for everyone
Superstars
Musiclab
References

Contagion

Definition:

- (1) The spreading of a quality or quantity between individuals in a population.
- (2) A disease itself:
  the plague, a blight, the dreaded lurgi, ...

Two main classes of contagion:

1. Infectious diseases:
   tuberculosis, HIV, ebola, SARS, influenza, ...
2. Social contagion:
   fashion, word usage, rumors, riots, religion, ...

Contagion models

Some large questions concerning network contagion:

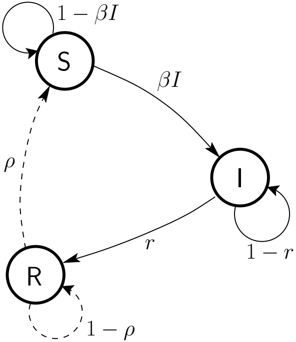
1. For a given spreading mechanism on a given network, what's the probability that there will be global spreading?
2. If spreading does take off, how far will it go?
3. How do the details of the network affect the outcome?
4. How do the details of the spreading mechanism affect the outcome?
5. What if the seed is one or many nodes?

Mathematical Epidemiology

The standard SIR model:

- Three states:
  - S = Susceptible
  - I = Infected
  - R = Recovered
- $S(t) + I(t) + R(t) = 1$
- Presumes random interactions

Discrete time example:



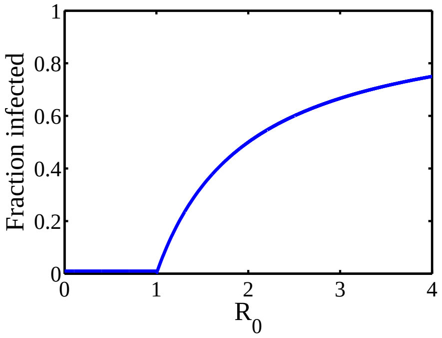
Transition Probabilities:

β for being infected given contact with infected
r for recovery
ρ for loss of immunity

Independent Interaction models

Reproduction Number R_0:

- R_0 = expected number of infected individuals resulting from a single initial infective.
- Epidemic threshold: If $R_0 > 1$, 'epidemic' occurs.
- Example:



- Continuous phase transition.
- Fine idea from a simple model.

Disease spreading models

For 'novel' diseases:

1. Can we predict the size of an epidemic?
2. How important/useful is the reproduction number R_0?
3. What is the population size N?

R_0 and variation in epidemic sizes

R_0 approximately the same for all of the following:

- 1918-19 “Spanish Flu” $\sim$ 500,000 deaths in US
- 1957-58 “Asian Flu” $\sim$ 70,000 deaths in US
- 1968-69 “Hong Kong Flu” $\sim$ 34,000 deaths in US
- 2003 “SARS Epidemic” $\sim$ 800 deaths world-wide

Size distributions

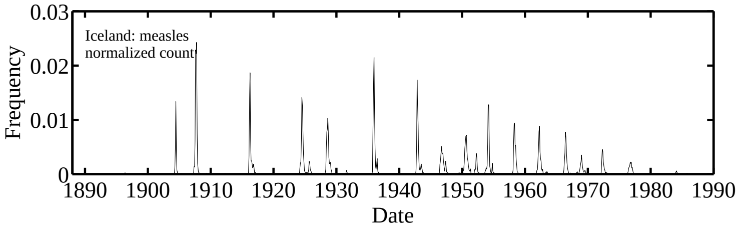
Elsewhere, event size distributions are important:

- earthquakes (Gutenberg-Richter law)
- city sizes, forest fires, war fatalities
- wealth distributions
- 'popularity' (books, music, websites, ideas)
- What about Epidemics?

Power laws distributions are common but not obligatory...

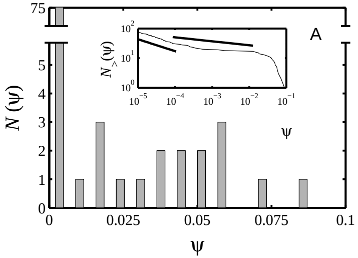
Feeling icky in Iceland

Caseload recorded monthly for range of diseases in Iceland, 1888-1990



Treat outbreaks separated in time as ‘novel’ diseases.

Measles



Insert plots:

Complementary cumulative frequency distributions:

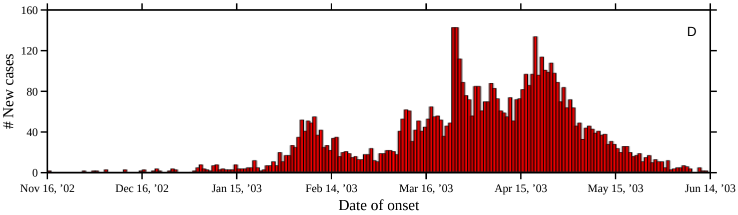
$$N_>(\Psi) \propto \Psi^{-\gamma+1}$$

Ψ = fractional epidemic size

Measured values of γ:

- measles: 1.40 (low Ψ) and 1.13 (high Ψ)
- Expect $2 \leq \gamma < 3$ (finite mean, infinite variance)
- Distribution is rather flat...

Resurgence—example of SARS



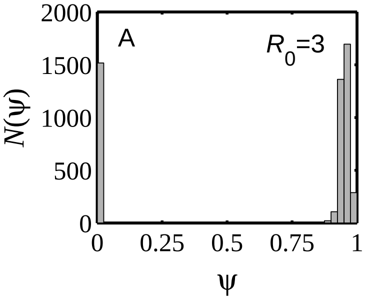
- Epidemic discovers new 'pools' of susceptibles: Resurgence.
- Importance of rare, stochastic events.

A challenge

So... can a simple model produce

1. broad epidemic distributions
   and
2. resurgence ?

Size distributions

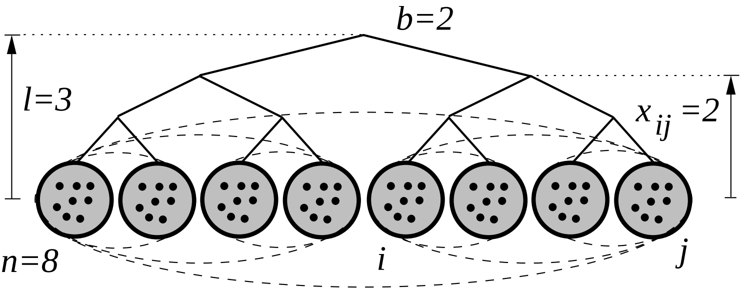


Simple models typically produce bimodal or unimodal size distributions.

- This includes network models: random, small-world, scale-free, ...
- Some exceptions:
  1. Forest fire models
  2. Sophisticated metapopulation models

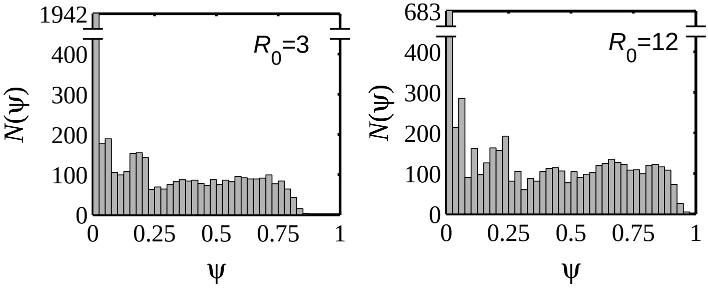
A toy agent-based model

Geography: allow people to move between contexts:



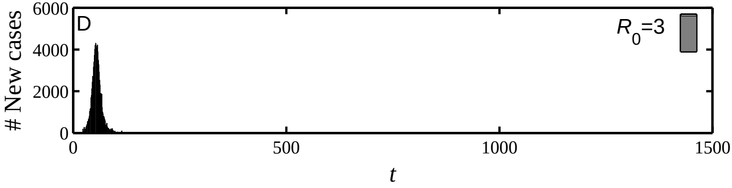
- P = probability of travel
- Movement distance: $\Pr(d) \propto \exp(-d/\xi)$
- ξ = typical travel distance

Example model output: size distributions

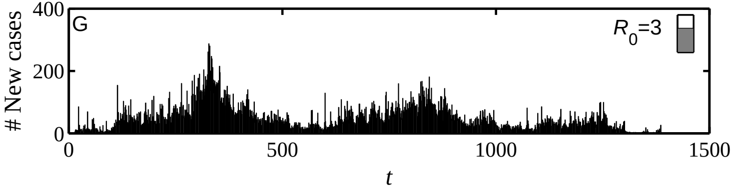


- Flat distributions are possible for certain ξ and P.
- Different R_0's may produce similar distributions
- Same epidemic sizes may arise from different R_0's

Standard model:



Standard model with transport: Resurgence



- Disease spread highly sensitive to population structure
- Rare events may matter enormously

Simple disease spreading models

Attempts to use beyond disease:

- Adoption of ideas/beliefs (Goffman & Newell, 1964)
- Spread of rumors (Daley & Kendall, 1965)
- Diffusion of innovations (Bass, 1969)
- Spread of fanatical behavior (Castillo-Chávez & Song, 2003)

Social Contagion

Social Contagion

Examples abound:

- being polite/rude
- strikes
- innovation
- residential segregation
- ipods
- obesity
- Harry Potter
- voting
- gossip
- Rubik's cube 
- religious beliefs
- leaving lectures

SIR and SIRS contagion possible

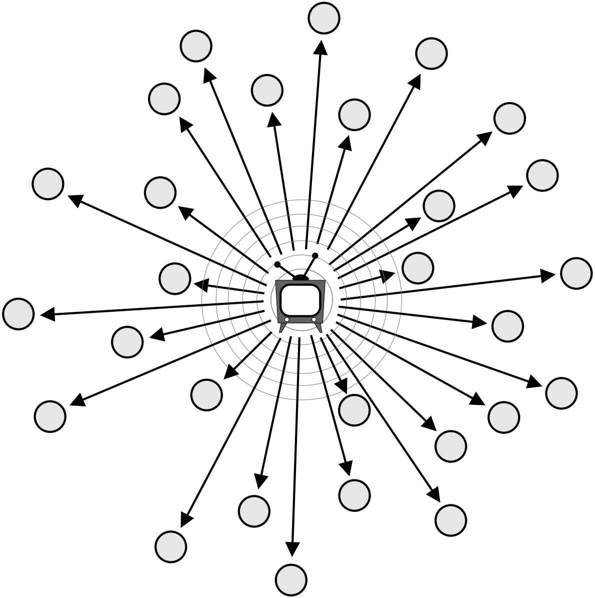
- Classes of behavior versus specific behavior: dieting

Social Contagion

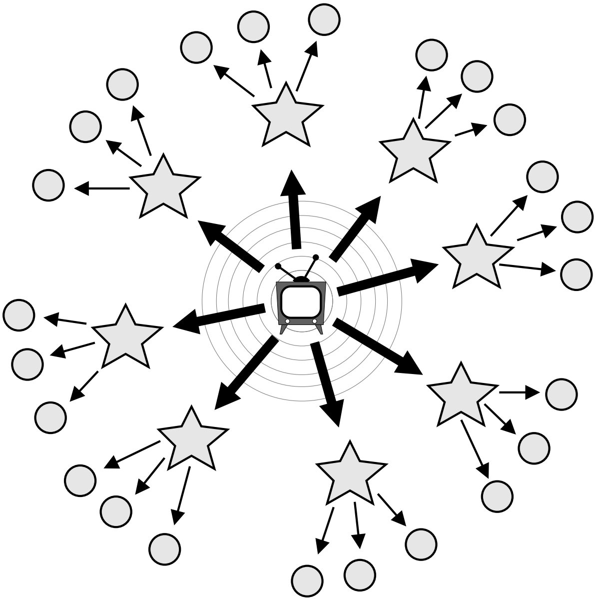
Two focuses for us:

- Widespread media influence
- Word-of-mouth influence

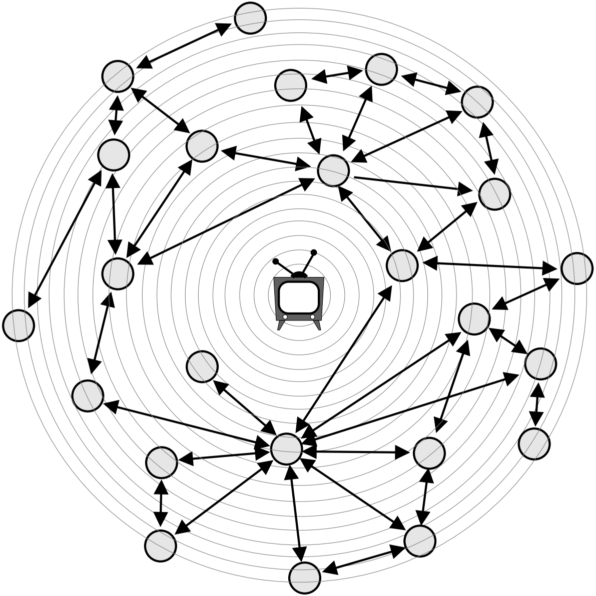
The hypodermic model of influence:



The two step model of influence:



The general model of influence:



Social Contagion

Why do things spread?

- Because of system level properties?
- Or properties of special individuals?
- Is the match that lights the forest fire the key? (Katz and Lazarsfeld; Gladwell)
- Yes. But only because we are narrative-making machines...
- System/group properties harder to understand
- Always good to examine what is said before and after the fact...

The Mona Lisa:

- "Becoming Mona Lisa: The Making of a Global Icon"—David Sassoon
- Not the world's greatest painting from the start...
- Escalation through theft, vandalism, parody, ...

The completely unpredicted fall of Eastern Europe:

Timur Kuran: “Now Out of Never: The Element of Surprise in the East European Revolution of 1989”

Social Contagion

Some important models:

- Tipping models—Schelling (1971)
  - Simulation on checker boards
  - Idea of thresholds
- Threshold models—Granovetter (1978)
- Herding models—Bikhchandani, Hirschleifer, Welch (1992)
  - Social learning theory, Informational cascades,...

Social contagion models

Thresholds:

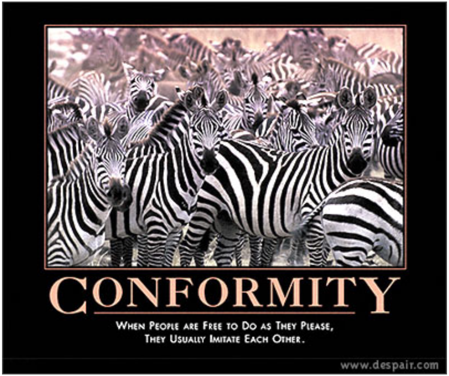
- Basic idea: individuals adopt a behavior when a certain fraction of others have adopted
- 'Others' may be everyone in a population, an individual's close friends, any reference group.
- Response can be probabilistic or deterministic.
- Individual thresholds vary.

Social Contagion

Some possible origins of thresholds:

- Desire to coordinate, to conform.
- Lack of information: impute the worth of a good or behavior based on degree of adoption (social proof)
- Economics: Network effects or network externalities
  - Telephones, Facebook, operating systems, ...

Imitation



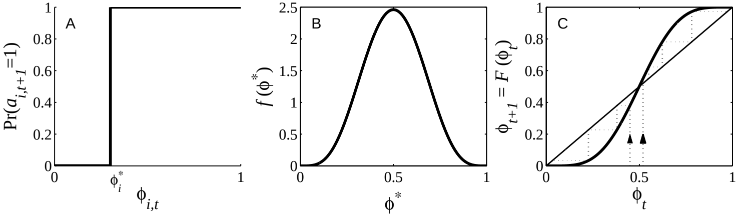
despair.com

"When people are free to do as they please, they usually imitate each other."

—Eric Hoffer
"The Passionate State of Mind"[11]

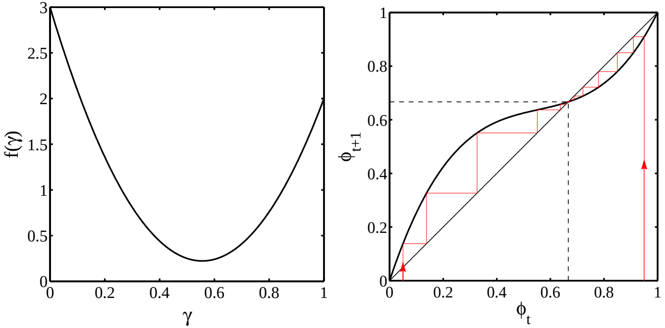
Granovetter's threshold model:

Action based on perceived behavior of others:



- Two states: S and I.
- ϕ = fraction of contacts 'on' (e.g., rioting)
- $$\phi_{t+1} = \int_0^{\phi_t} f(\gamma)\mathrm{d}\gamma = F(\gamma)|_0^{\phi_t} = F(\phi_t)$$
- This is a Critical Mass model

Social Sciences: Threshold models



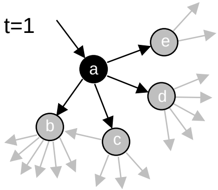
- Example of single stable state model

Social Sciences—Threshold models

Implications for collective action theory:

1. Collective uniformity $\not\Rightarrow$ individual uniformity
2. Small individual changes $\Rightarrow$ large global changes

Threshold model on a network



- All nodes have threshold $\phi = 0.2$.
- “A simple model of global cascades on random networks”
  D. J. Watts. Proc. Natl. Acad. Sci., 2002

Snowballing

The Cascade Condition:

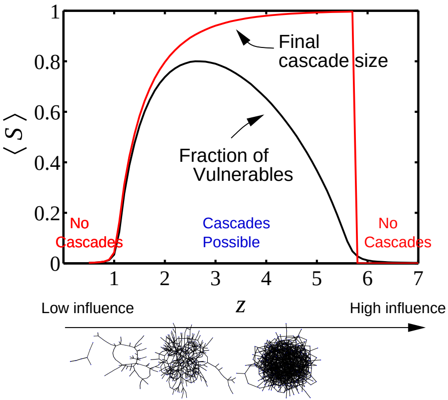
- If one individual is initially activated, what is the probability that an activation will spread over a network?
- What features of a network determine whether a cascade will occur or not?

The most gullible

Vulnerables:

- = Individuals who can be activated by just one 'infected' contact
- For global cascades on random networks, must have a *global cluster of vulnerables*
- Cluster of vulnerables = critical mass
- Network story: 1 node $\rightarrow$ critical mass $\rightarrow$ everyone.

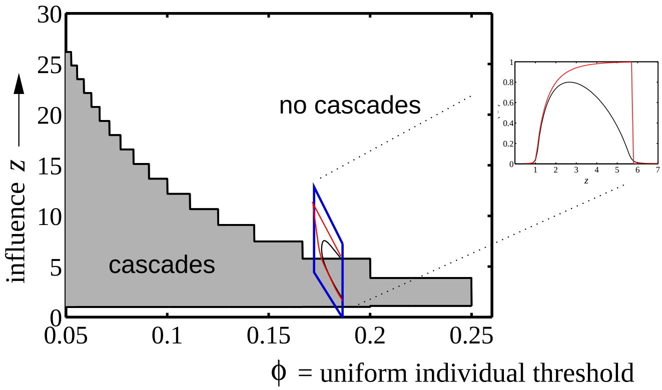
Cascades on random networks



Example networks

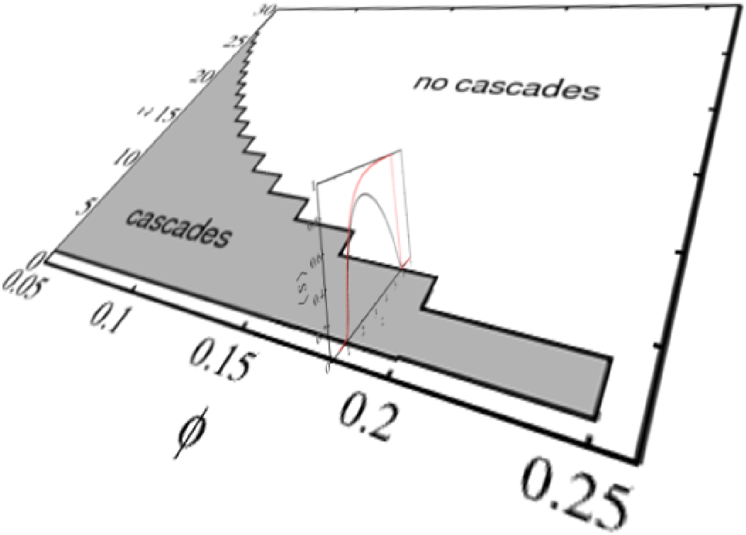
- Cascades occur only if size of max vulnerable cluster > 0.
- System may be 'robust-yet-fragile'.
- 'Ignorance' facilitates spreading.

Cascade window for random networks



- 'Cascade window' widens as threshold ϕ decreases.
- Lower thresholds enable spreading.

Cascade window for random networks



Analytic work

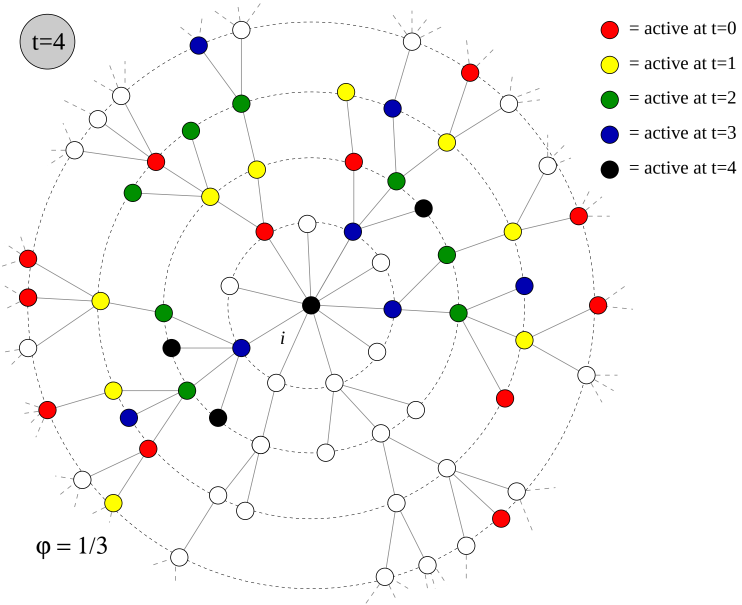
- Threshold model completely solved (by 2008):
- Cascade condition:[22]

$$\sum_{k=1}^{\infty} k(k-1)\beta_k P_k / z \geq 1.$$

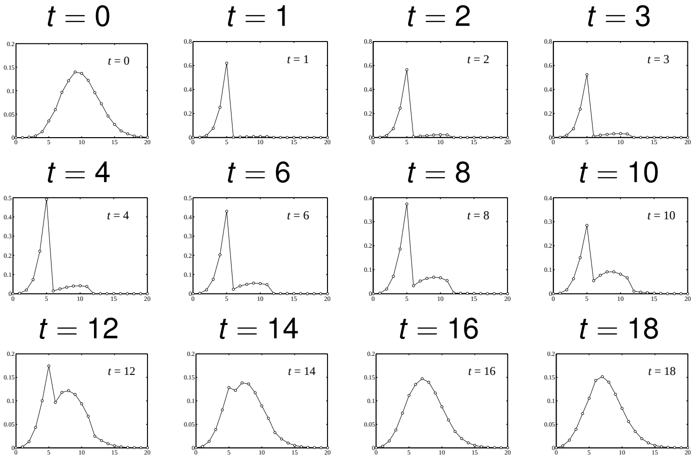
where β_k = probability a degree k node is vulnerable.

- Final size of spread figured out by Gleeson and Calahane[9, 8].
- Solution involves finding fixed points of an iterative map of the interval.
- Spreading takes off: expansion
- Spreading reaches a particular node: contraction

Expected size of spread

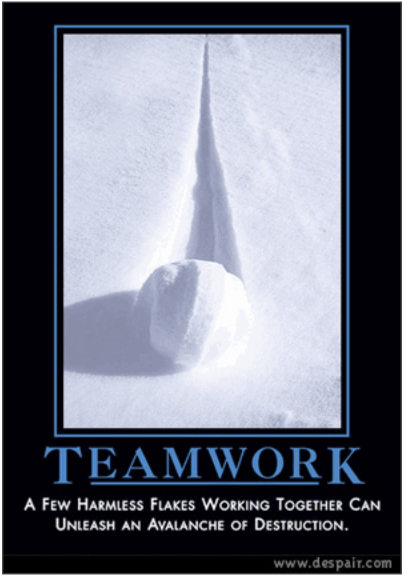


Early adopters—degree distributions



$P_{k,t}$ versus k

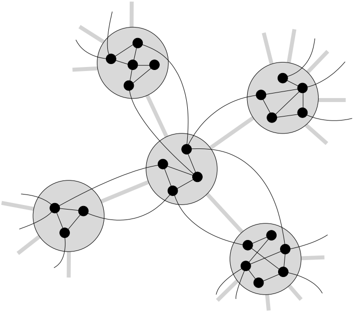
The power of groups...



despair.com

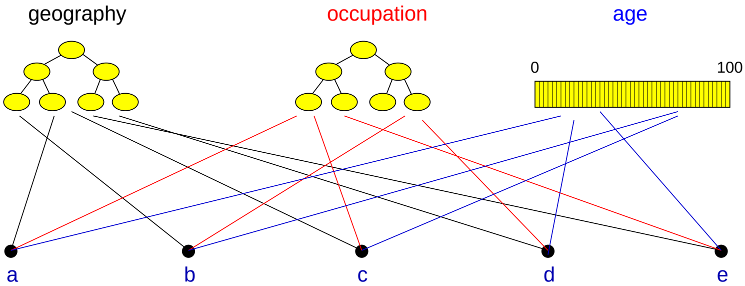
"A few harmless flakes working together can unleash an avalanche of destruction."

Group structure—Ramified random networks



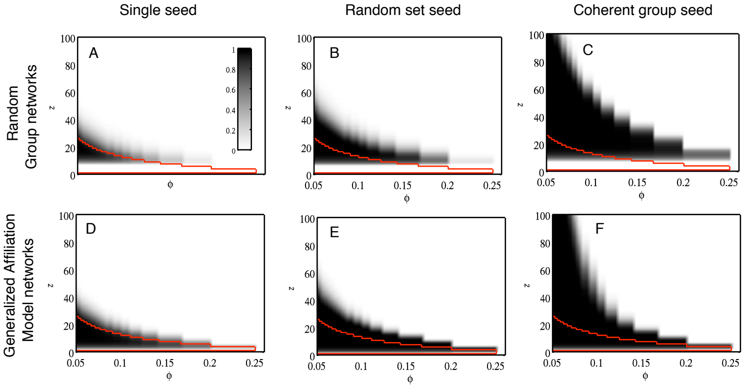
p = intergroup connection probability
q = intragroup connection probability.

Generalized affiliation model

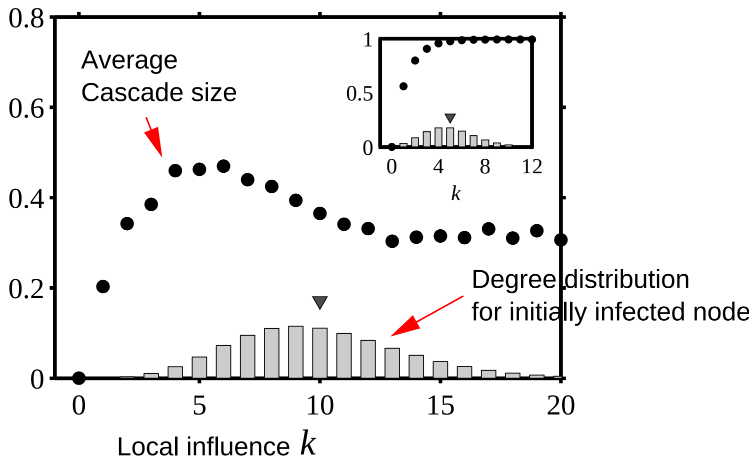


(Blau & Schwartz, Simmel, Breiger)

Cascade windows for group-based networks



Assortativity in group-based networks



- The most connected nodes aren't always the most 'influential.'
- Degree assortativity is the reason.

Social contagion

Summary:

- 'Influential vulnerables' are key to spread.
- Early adopters are mostly vulnerables.
- Vulnerable nodes important but not necessary.
- Groups may greatly facilitate spread.
- Extreme/unexpected cascades may occur in highly connected networks
- Many potential 'influentials' exist.
- Average individuals may be more influential system-wise than locally influential individuals.
- 'Influentials' are posterior constructs.

Social contagion

Implications:

- Focus on the influential vulnerables.
- Create entities that many individuals 'out in the wild' will adopt and display rather than broadcast from a few 'influentials.'
- Displaying can be passive = free (yo-yo's, fashion), or active = harder to achieve (political messages).
- Accept that movement of entities will be out of originator's control.
- Possibly only simple ideas can spread by word-of-mouth.

  (Idea of opinion leaders has spread well...)

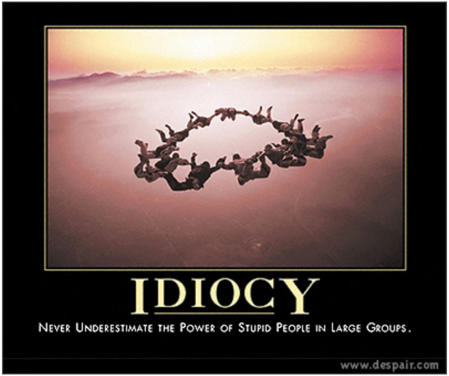
Social Contagion

Messing with social connections:

- Ads based on message content (e.g., Google and email)
- Buzz media
- Facebook's advertising (Beacon)

Arguably not always a good idea...

The collective...



despair.com

"Never Underestimate the Power of Stupid People in Large Groups."

Where do superstars come from?

Rosen (1981): "The Economics of Superstars"

Examples:

- Full-time Comedians ($\approx$ 200)
- Soloists in Classical Music
- Economic Textbooks (the usual myopic example)

- Highly skewed distributions again...

Superstars

Rosen's theory:

- Individual quality q maps to reward $R(q)$
- $R(q)$ is 'convex' ($\mathrm{d}^2 R/\mathrm{d}q^2 > 0$)
- Two reasons:
  1. Imperfect substitution:
     A very good surgeon is worth many mediocre ones
  2. Technology:
     Media spreads & technology reduces cost of reproduction of books, songs, etc.
- No social element—success follows 'inherent quality'

Superstars

Adler (1985): "Stardom and Talent"

- Assumes extreme case of equal 'inherent quality'
- Argues desire for coordination in knowledge and culture leads to differential success
- Success is then purely a social construction

Dominance hierarchies

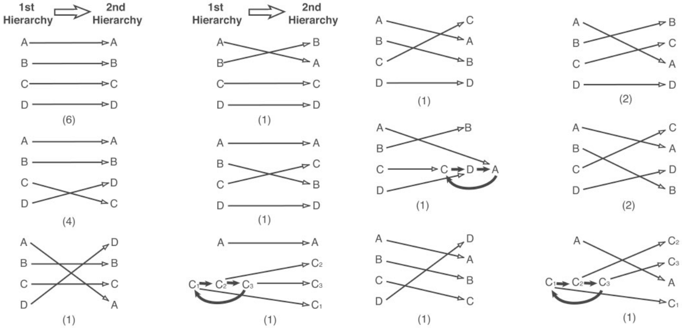
Chase et al. (2002): "Individual differences versus social dynamics in the formation of animal dominance hierarchies"

The aggressive female Metriaclima zebra (⊞):

Pecking orders for fish...

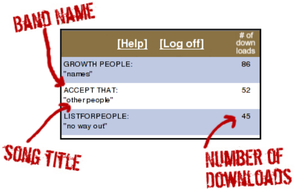
Dominance hierarchies

- Fish forget—changing of dominance hierarchies:



- 22 observations: about 3/4 of the time, hierarchy changed

Music Lab Experiment



48 songs
30,000 participants

multiple 'worlds'
Inter-world variability

- How probable is the world?
- Can we estimate variability?
- Superstars dominate but are unpredictable. Why?

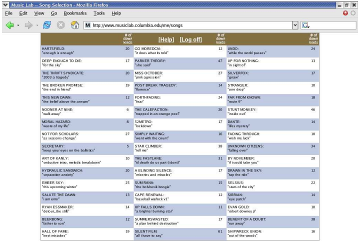
Music Lab Experiment

	# of down loads		# of down loads		# of down loads
HARTSFIELD: "enough is enough"	20	GO MOREDCAI: "it does what its told"	12	UNDO: "while the world passes"	24
DEEP ENOUGH TO DIE: "for the sky"	17	PARKER THEORY: "she said"	47	UP FOR NOTHING: "in sight of"	13
THE THRIFT SYNDICATE: "2003 a tragedy"	20	MISS OCTOBER: "pink agression"	27	SILVERFOX: "gnaw"	17
THE BROKEN PROMISE: "the end in friend"	19	POST BREAK TRAGEDY: "florence"	14	STRANGER: "one drop"	10
THIS NEW DAWN: "the belief above the answer"	12	FORTHFADING: "fear"	24	FAR FROM KNOWN: "route 9"	18
NOONER AT NINE: "walk away"	6	THE CALEFACTION: "trapped in an orange peel"	20	STUNT MONKEY: "inside out"	46
MORAL HAZARD: "waste of my life"	8	52METRO: "lockdown"	17	DANTE: "lifes mystery"	14
NOT FOR SCHOLARS: "as seasons change"	27	SIMPLY WAITING: "went with the count"	16	FADING THROUGH: "wish me luck"	10
SECRETARY: "keep your eyes on the ballistics"	5	STAR CLIMBER: "tell me"	38	UNKNOWN CITIZENS: "falling over"	34
ART OF KANLY: "seductive intro, melodic breakdown"	10	THE FASTLANE: "til death do us part (i dont)"	31	BY NOVEMBER: "if i could take you"	20
HYDRAULIC SANDWICH: "separation anxiety"	20	A BLINDING SILENCE: "miseries and miracles"	17	DRAWN IN THE SKY: "tap the ride"	12
EMBER SKY: "this upcoming winter"	25	SUM RANA: "the bolshevik boogie"	15	SELSIUS: "stars of the city"	22
SALUTE THE DAWN: "i am error"	13	CAPE RENEWAL: "baseball warlock v1"	12	SIBRIAN: "eye patch"	14
RYAN ESSMAKER: "detour_(be still)"	14	UP FALLS DOWN: "a brighter burning star"	11	EVAN GOLD: "robert downey jr"	10
BEERBONG: "father to son"	12	SUMMERSWASTED: "a plan behind destruction"	17	BENEFIT OF A DOUBT: "run away"	38
HALL OF FAME: "best mistakes"	19	SILENT FILM: "all i have to say"	61	SHIPWRECK UNION: "out of the woods"	16

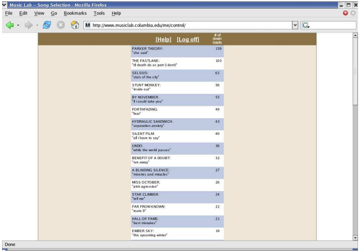
Salganik et al. (2006) "An experimental study of inequality and unpredictability in an artificial cultural market"

Music Lab Experiment

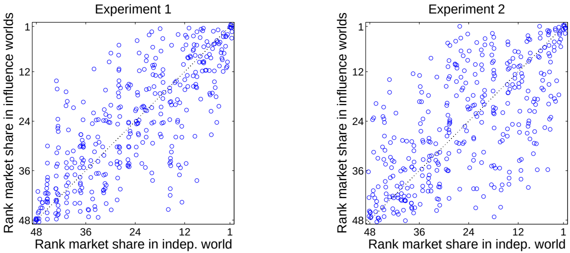
Experiment 1



Experiments 2–4

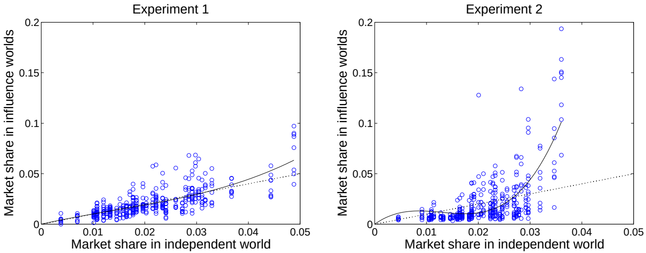


Music Lab Experiment



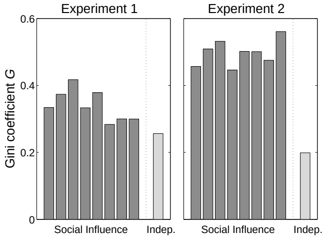
- Variability in final rank.

Music Lab Experiment



- Variability in final number of downloads.

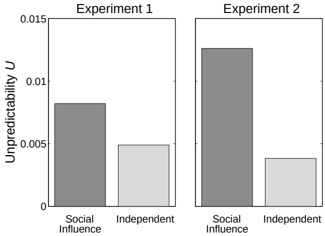
Music Lab Experiment



- Inequality as measured by Gini coefficient:

$$G = \frac{1}{(2N_s - 1)} \sum_{i=1}^{N_s} \sum_{j=1}^{N_s} |m_i - m_j|$$

Music Lab Experiment



- Unpredictability

$$U = \frac{1}{N_s \binom{N_w}{2}} \sum_{i=1}^{N_s} \sum_{j=1}^{N_w} \sum_{k=j+1}^{N_w} |m_{i,j} - m_{i,k}|$$

Music Lab Experiment

Sensible result:

- Stronger social signal leads to greater following and greater inequality.

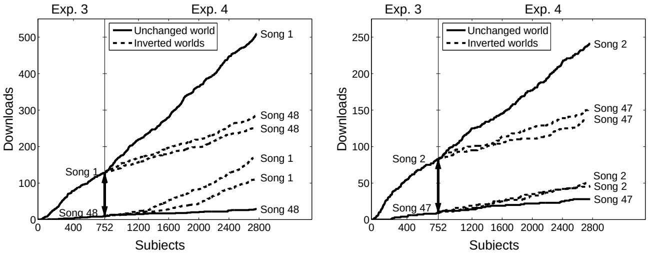
Peculiar result:

- Stronger social signal leads to greater unpredictability.

Very peculiar observation:

- The most unequal distributions would suggest the greatest variation in underlying ‘quality.’
- But success may be due to social construction through following...

Music Lab Experiment—Sneakiness



- Inversion of download count
- The 'pretend rich' get richer ...
- ... but at a slower rate

References I

[1] M. Adler.
Stardom and talent.
American Economic Review, pages 208–212, 1985.
pdf (⊞)

[2] S. Bikhchandani, D. Hirshleifer, and I. Welch.
A theory of fads, fashion, custom, and cultural change as informational cascades.
J. Polit. Econ., 100:992–1026, 1992.

[3] S. Bikhchandani, D. Hirshleifer, and I. Welch.
Learning from the behavior of others: Conformity, fads, and informational cascades.
J. Econ. Perspect., 12(3):151–170, 1998. pdf (⊞)

References II

[4] J. Carlson and J. Doyle.
Highly optimized tolerance: A mechanism for power laws in design systems.
Phys. Rev. Lett., 60(2):1412–1427, 1999. pdf (⊞)

[5] J. Carlson and J. Doyle.
Highly optimized tolerance: Robustness and design in complex systems.
Phys. Rev. Lett., 84(11):2529–2532, 2000. pdf (⊞)

[6] I. D. Chase, C. Tovey, D. Spangler-Martin, and M. Manfredonia.
Individual differences versus social dynamics in the formation of animal dominance hierarchies.
Proc. Natl. Acad. Sci., 99(8):5744–5749, 2002. pdf (⊞)

References III

[7] M. Gladwell.
The Tipping Point.
Little, Brown and Company, New York, 2000.

[8] J. P. Gleeson.
Cascades on correlated and modular random networks.
Phys. Rev. E, 77:046117, 2008. pdf (⊞)

[9] J. P. Gleeson and D. J. Cahalane.
Seed size strongly affects cascades on random networks.
Phys. Rev. E, 75:056103, 2007. pdf (⊞)

[10] M. Granovetter.
Threshold models of collective behavior.
Am. J. Sociol., 83(6):1420–1443, 1978. pdf (⊞)

References IV

[11] E. Hoffer.
The Passionate State of Mind: And Other Aphorisms.
Buccaneer Books, 1954.

[12] E. Katz and P. F. Lazarsfeld.
Personal Influence.
The Free Press, New York, 1955.

[13] T. Kuran.
Now out of never: The element of surprise in the east european revolution of 1989.
World Politics, 44:7–48, 1991.

[14] T. Kuran.
Private Truths, Public Lies: The Social Consequences of Preference Falsification.
Harvard University Press, Cambridge, MA, Reprint edition, 1997.

References V

[15] J. D. Murray.
Mathematical Biology.
Springer, New York, Third edition, 2002.

[16] S. Rosen.
The economics of superstars.
Am. Econ. Rev., 71:845–858, 1981. pdf (⊞)

[17] M. J. Salganik, P. S. Dodds, and D. J. Watts.
An experimental study of inequality and unpredictability in an artificial cultural market.
Science, 311:854–856, 2006. pdf (⊞)

[18] T. Schelling.
Dynamic models of segregation.
J. Math. Sociol., 1:143–186, 1971.

References VI

[19] T. C. Schelling.
Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities.
J. Conflict Resolut., 17:381–428, 1973.

[20] T. C. Schelling.
Micromotives and Macrobehavior.
Norton, New York, 1978.

[21] D. Sornette.
Critical Phenomena in Natural Sciences.
Springer-Verlag, Berlin, 2nd edition, 2003.

[22] D. J. Watts.
A simple model of global cascades on random networks.
Proc. Natl. Acad. Sci., 99(9):5766–5771, 2002.
pdf (⊞)