

Information and Influence Spread in Social Networks

Carlos Castillo (Qatar Computing Research Institute)

Wei Chen (Microsoft Research Asia)

Laks V. S. Lakshmanan* (University of British Columbia)

* Research supported by an NSERC Strategic
Grant on Business Intelligence Network



Disclaimers

- What this tutorial will not cover/do
 - Comprehensive Study of various Diffusion Models, Applications, Network Measurements.
 - General Graph Mining
 - Analytics of Social Media
 - Heterogeneous Information Networks
 - No completeness guarantee on focal topic!
- Where to look if you are interested in these topics?
 - Information Diffusion and Influentials [Budak, Al Abbadi, and Agrawal VLDB 2011]
 - Graph Mining [Faloutsos, Miller and Tsourakakis KDD 2009]
 - Social Media Analytics [Leskovec KDD 2011]
 - Heterogeneous IN [Han, Sun, Yan, and Yu, KDD 2010].

Acknowledgments



Francesco Bonchi
Yahoo! Research



Amit Goyal
PhD Student
UBC



Michalis
Mathioudakis
Phd Student
Univ. Toronto



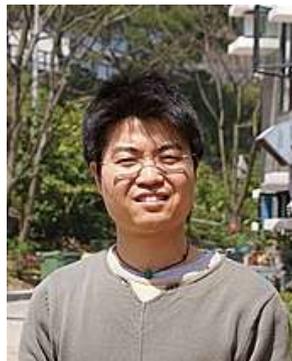
Smriti Bhagat
Technicolor Palo
Alto Research Lab



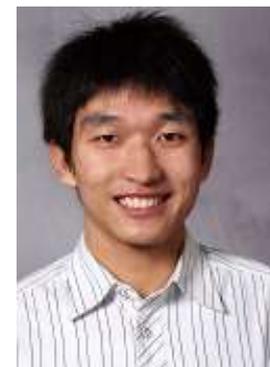
Suresh Venkata-
Subramanian
Utah



Wei Lu
PhD Student
UBC



Yajun Wang
MSRA



Chi Wang
PhD Student
UIUC

Acknowledgments (cont'd)



Chin-Yew Lin
MSRA



Li Zhang
MSR-SVC



Zhenming Liu
post-doc
Princeton



Guojie Song
Peking U.



Tao Sun
PhD student
Peking U.



Xiaorui Sun
PhD student
Columbia



Wei Wei
PhD student
CMU



Rachel Cummings
PhD student
Northwestern U.

Acknowledgments (cont'd)



Yifei Yuan
PhD student
UPenn



Xinran He
PhD student
USC



Te (Tony) Ke
PhD student
UC Berkeley



Alex Collins
Google



Siyu Yang
PhD student
Princeton



David Rincon
Universitat Politècnica
de Catalunya



Qingye Jiang
Master student
Columbia U.

Ning Zhang
PhD student
Purdue U.

Ming Zhang
Peking U.

Outline

- Part I: Motivation, Applications and Key Concepts
- Part II: Data and Tools
- Part III: Influence Maximization
- Part IV: Other Issues
- Part V: Challenges

Part I → Part II → Part III → Part IV → Part V

Motivation, Applications and Key Concepts



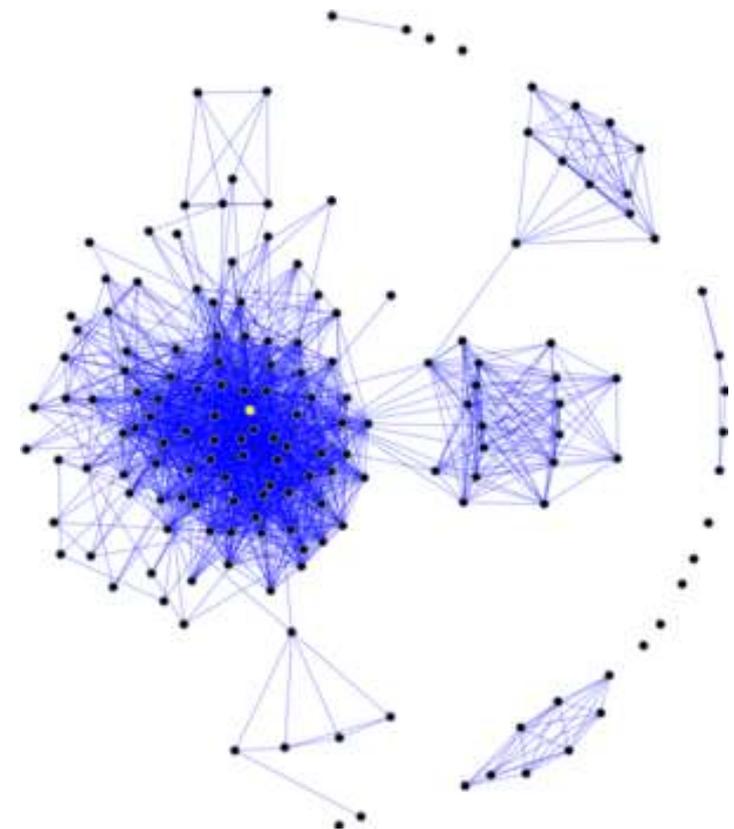
Part I: Outline

- Social Networks and Social Influence
- Real-world stories
- Example applications
- The Flip Side

Social Networks and Social Influence



Online Social Networking Sites



Social Networks & Media



SarcasticRover

@SarcasticRover

*Not the real @marscuriosity...
like I care.*

4th Rock From the Sun · <http://mars.jpl.nasa.gov/msl/>

 Follow

291 TWEETS

257 FOLLOWING

63,197 FOLLOWERS



Oh sure, I can't think of anything I'd rather be doing than driving around a wasteland looking at dirt for the rest of my life.

 **SarcasticRover**

3 days ago

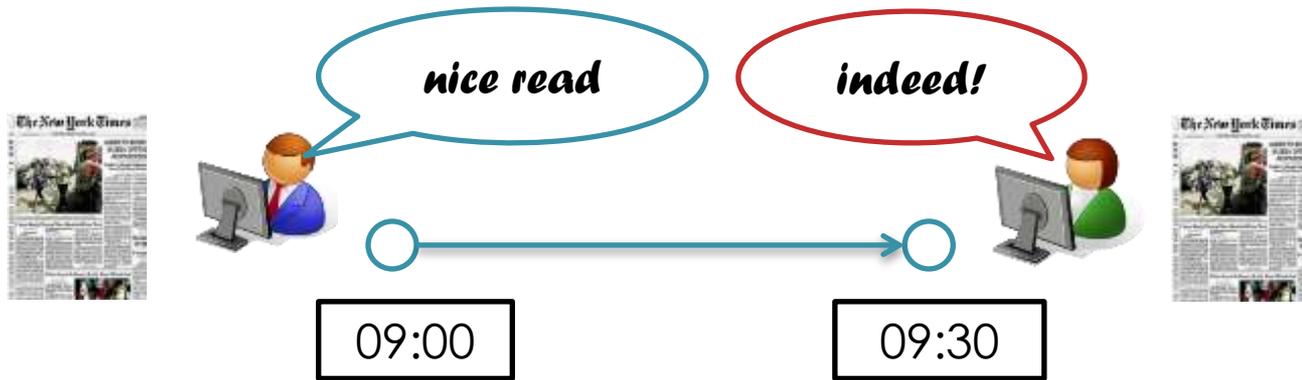


I'm really glad all you fricken hipsters took a vague interest in science for 8 hours. Thanks for that.

 **SarcasticRover**

3 days ago

Information Propagation



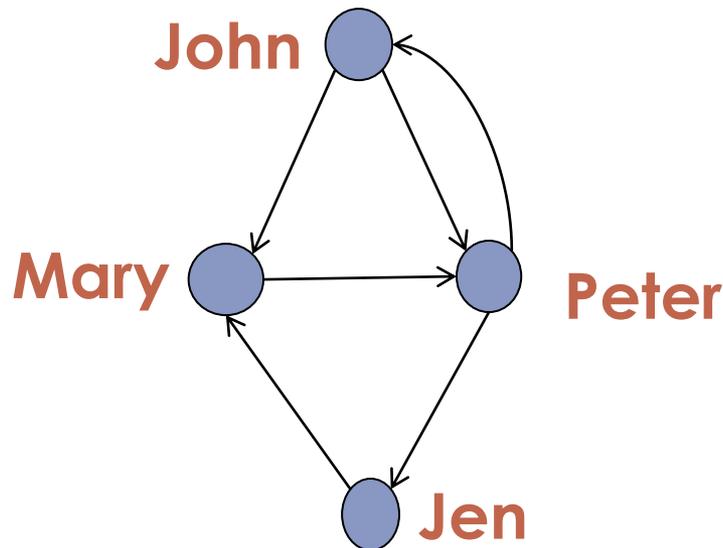
People are **connected** and perform **actions**

↓
friends, fans,
followers, etc.

↓
comment, link, rate, like,
retweet, post a message,
photo, or video, etc.

Basic Data Model

Graph: users, links/ties
 $G = (V, E)$



Log: user, action, time
 $A = \{ \langle u_1, a_1, t_1 \rangle, \dots \}$

User	Action	Time
John	Rates with 5 stars "The Artist"	June 3 rd
Peter	Watches "The Artist"	June 5 th
Jen

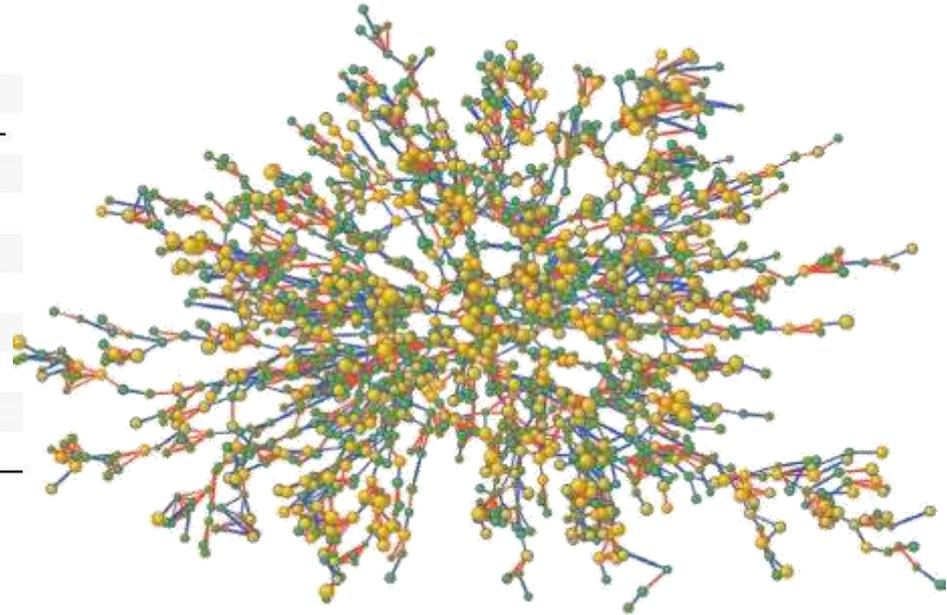
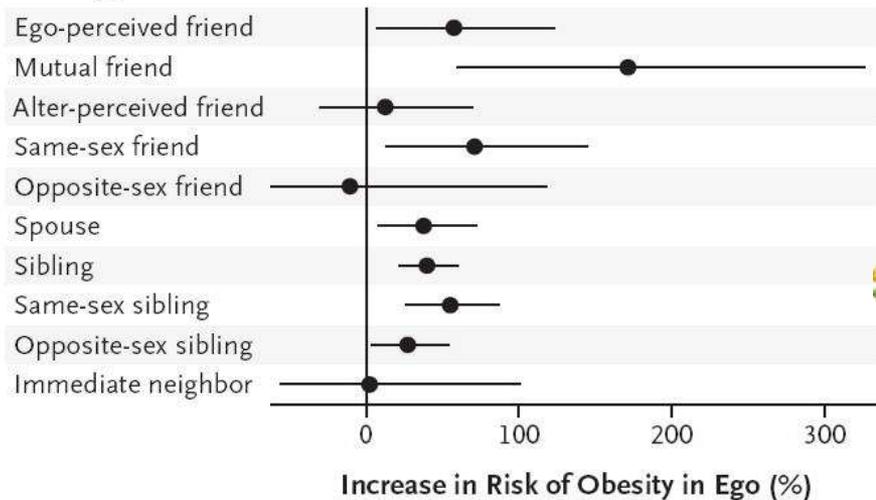
Real World Stories



Social Influence: Real-world Story I

12K people, 50K links, medical records from 1971 to 2003

Alter Type

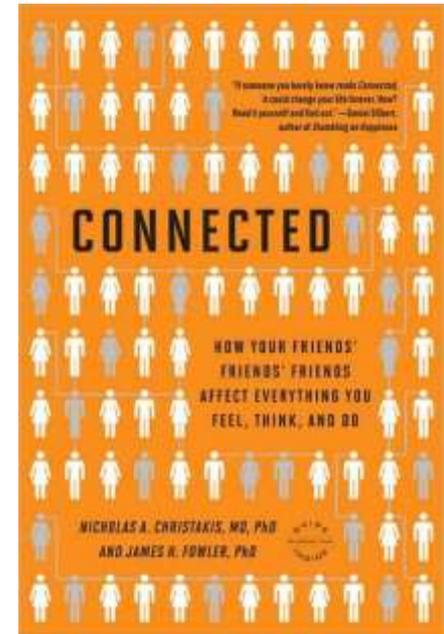


- Obese Friend → 57% increase in chances of obesity
- Obese Sibling → 40% increase in chances of obesity
- Obese Spouse → 37% increase in chances of obesity

Social Influence: Real-world Story II

Key to understanding people is understanding ties between them.

Your friend's friends' actions and feelings affect your thoughts, feelings and actions!



- **Back pain:** spread from West to East in Germany after fall of Berlin Wall
- **Suicide:** well known to spread throughout communities on occasion
- **Sex practices:** e.g., growing prevalence of oral sex among teenagers
- **Politics:** the denser your connections, the more intense your convictions

Social Media “Friends”

I would delete you off of my
facebook friends list, but
then you wouldn't be
able to see all the fun
things I do without
you.



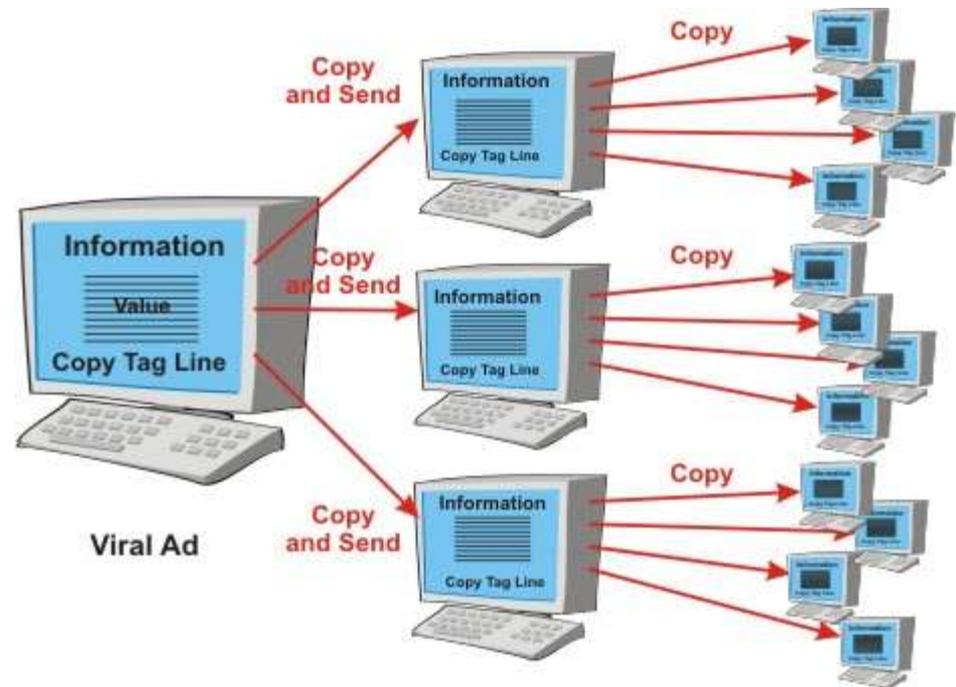
your  cards
someecards.com

Social Influence: Real-world Story III

- Hotmail's viral climb to the top spot (**90s**): 8 million users in 18 months!
- Boosted brand awareness
- Far more effective than conventional advertising by rivals
 - ... and far cheaper, too!

Join the world's largest e-mail service with MSN Hotmail. <http://www.hotmail.com>

Simple message added to footer of every email message sent out



Social Influence: Real-world Story IV

- From rags to riches – **Ted Williams**

- Voice over artist
- **Homeless** and many a brush with the law.
- Found at a street corner in Columbus, OH in Jan 2011
- Interview posted in YouTube; **13 million views**
- Attracted numerous offers, including **jobs!**



Social Influence: Real-world Story V



- Gold award from YouTube for most hits; featured in Time, BBC News, News1130 ...
> 58×10^6 hits on YouTube as of June 2012

Social Influence: Real-world Story V

- Indian song from the sound track of the upcoming Tamil movie *Why this கொலைவெறி டி?* (*Why this kolaveri di?*)
 - Released on Nov. 16, 2011
 - Top trend on [Twitter](#) on Nov. 21 2011
- Within 1 week of release:
 - > 1.3×10^6 views on [YouTube](#)
 - > 10^6 “shares” on [Facebook](#)
- Reaches many non-Tamil speakers.

Info. Diffusion: Real-world Story VI

2008 Mumbai Terror Attacks

- **≈16 tweets/second** sent to Twitter via SMS
 - eyewitness accounts, pleas for blood donors...
- **Wikipedia page** up within minutes, with staggering amount of detail and extremely fast “live” updates
- **Metroblog** as a newswire service; **112 Flickr photos** by a journalist giving a firsthand account of aftermath
- **Google map** with main buildings involved in the attacks, with links to background and new stories!

Info. Diffusion: Real-world Story VII

2011 Stanley Cup Riots Vancouver



Young rioters bragging in social media: e.g., posing with (looted) Gucci bags in front of burning cars.

- Triggered widespread reactions of disgust
 - Turned into a way to mobilize clean-ups
- Over time, catch the rioters and publicly shame them on SM
- 100 hours VHS footage from 1994 riots vs. 5000 hours of 100 types of digital video
 - Need for sophisticated and efficient analytics

Example Applications



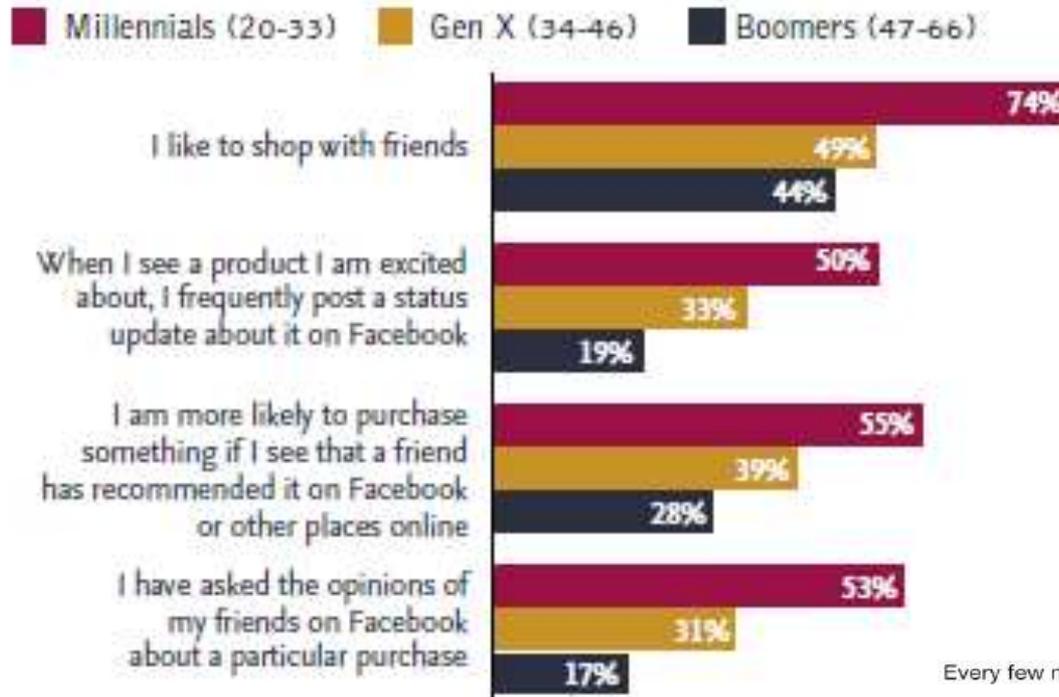
Applications

Viral Marketing

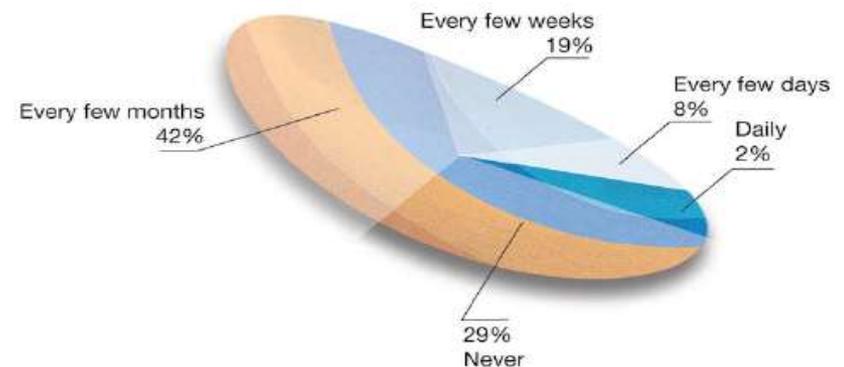
Social media analytics
Spread of falsehood and rumors
Interest, trust, referrals
Adoption of innovations
Human and animal epidemics
Expert finding
Behavioral targeting
Feed ranking
“Friends” recommendation
Social search

Application: viral marketing

Purchase decisions are increasingly influenced
by opinions of friends in Social Media



How frequently do you share recommendations online?



Viral/Word-of-Mouth Marketing

- Idea: **exploit social influence for marketing**
- Basic assumption: word-of-mouth effect
 - Actions, opinions, buying behaviors, innovations, etc. propagate in a social network
- **Target** users who are likely to produce word-of-mouth diffusion
 - Additional reach, clicks, conversions, brand awareness
 - **Target the influencers**



Transitivity of trust

- **Trust** is associated with the belief of an agent in the assertions by other agents; it is neither necessary nor sufficient for influence
- The **Web of Trust** from the early 1990s
 - Public Key Certification
 - *Advogato*: propagate trust through links
- **Transitive social importance** from the late 1940s
 - Seeley 1949, Wei 1952, Katz 1953: transitive importance computation
 - Reinvented as **PageRank** [Page et al. TR 1998]
 - TrustRank [Gyongyi et al. VLDB 2004], EigenTrust, Trust/distrust propagation

Social networks & marketing

SOCIAL MEDIA MARKETING MADNESS



Identifying influencers

- Influencers **increase** brand awareness.
product conversions through WoMM
 - Influencers **advocate** a brand
 - Influencers **influence** purchasing actions



Identifying influencers: start-ups

- **Klout**

- Measure of overall influence online (mostly Twitter, now FB and LinkedIn)
- Score = function of true reach, amplification probability and network influence
- Claims score to be highly correlated to clicks, comments and retweets

- **Peer Index**

- Identifies/Scores authorities on the social web by topic

- **SocialMatica**

- Ranks 32M people by vertical/topic, claims to take into account quality of authored content

- **Influencer50**

- Clients: IBM, Microsoft, SAP, Oracle and a long list of tech companies

+ Svnetwork, Bluecalypso, CrowdBooster, Sproutsocial, TwentyFeet, EmpireAvenue, Twitaholic, and many others ...

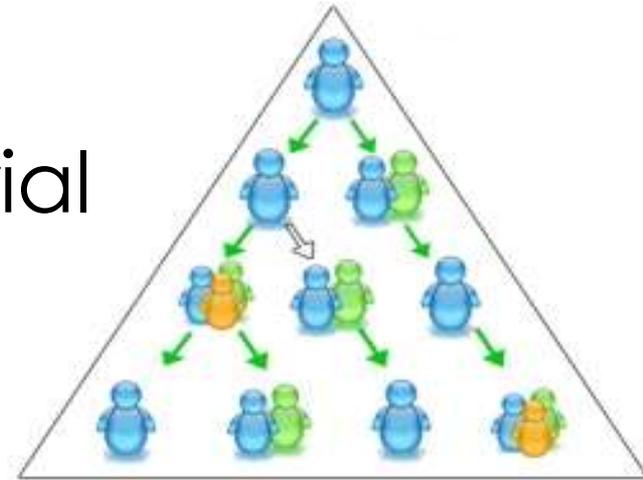
Finding the influencers ...



"He's not a 'Super Influencer',
he's a very naughty boy!"

Viral marketing & The Influence Maximization Problem

- Problem statement:
 - **find a seed-set of influential people such that by targeting them we maximize the spread of viral propagations**
- Focus of **Part III** of this tutorial



The Flip Side



Criticisms / caveats

1. Are we observing correlation or causation? Homophily or influence?
2. Can social influence actually drive viral cascades?
3. Is viral marketing useful in practice?



Homophily or Influence?

Homophily: tendency to stay together with people similar to you

“Birds of a feather flock together”

E.g. I'm overweight → I date overweight girls

Influence: force that a person A exerts on a person B that changes the behavior/opinion of B

Influence is a causal process

E.g. my girlfriend gains weight → I gain weight too

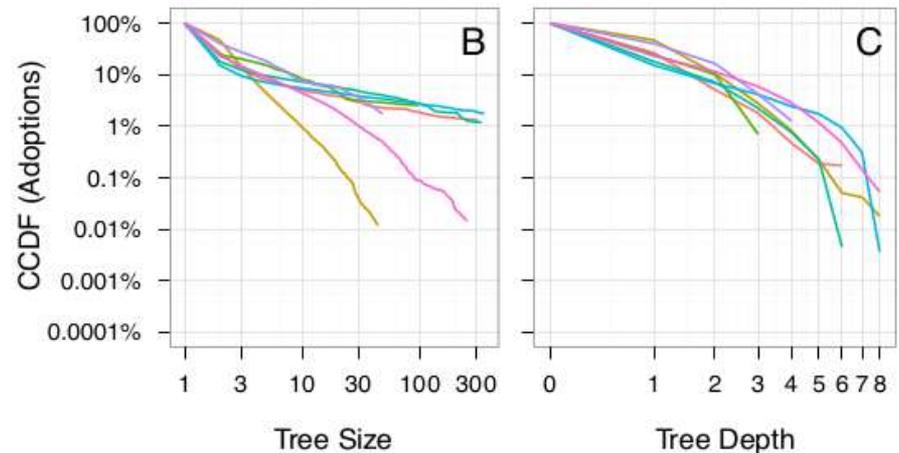
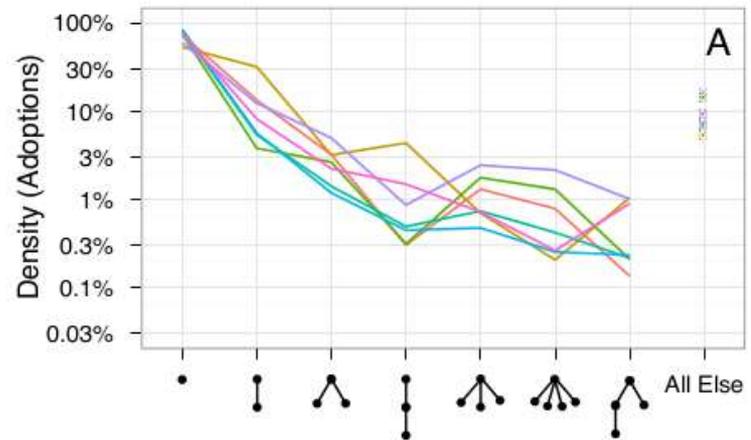
Can social influence really drive viral cascades?

- Watts et al. challenge the traditional notions and intuitions about SI causing viral spread
- Social epidemics are not always responsible for dramatic, possibly sudden social change
- Influence is hard to prove
- Do not dismiss influence altogether

“Viral” cascades are shallow

Across multiple social media platforms:

- Most adoptions are not due to influence from others (depth=0)
- Most cascades are shallow (depth=1-2)



[Goel et al. EC 2012]

How useful is viral marketing?

- Criticism #1: Hard to predict which campaign will succeed virally.
 - Lack of predictability makes VM hard to implement;
 - The magic might not be in a small number of influentials
 - “*Big seed*” marketing is a predictable, practical alternative

Example: Huffington Post

- Ad agency buys all of the ad slots for a week
- Displays attractive videos with options for easy sharing
- **Gets 7x more views** due to social referrals, but ...
- **None of the videos “goes viral”** (grows exponentially in views) at any time



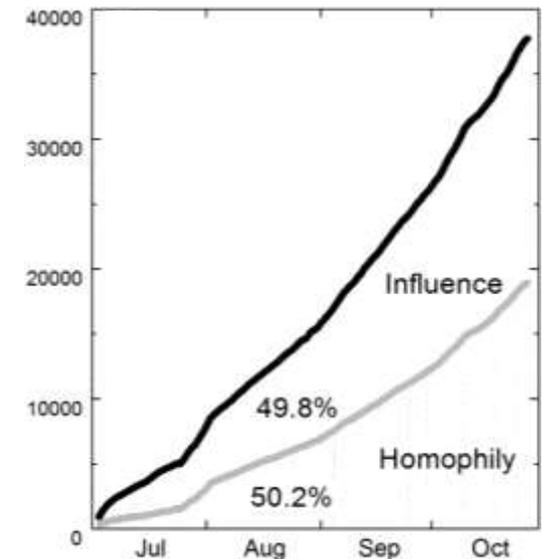
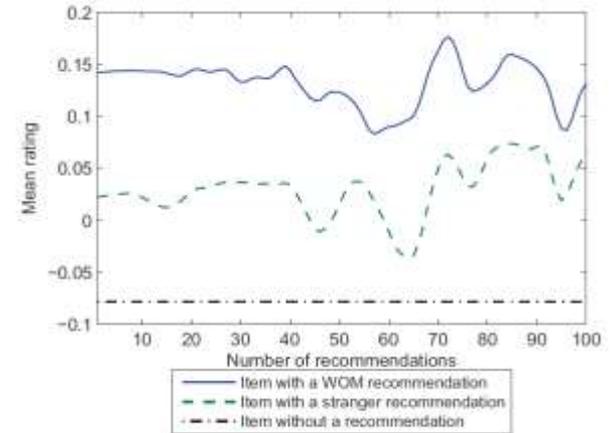
Watch "Gun Free"



Watch "Meth - David"

Evidence of Real Influence

- People rate an item higher if a friend has recommended it
- Even after carefully removing homophily effects, influence can be clearly observed



What Did We Learn So Far?



Key takeaways

- General idea of Social networks and information propagation and how they are modeled
- Several real-life stories of influence and information propagation
- Other applications
- The flip side: it is easy to get overexcited – both about existence of influence and about its absence!