

# Gesundheit!

## Modeling Contagion through Facebook News Feed

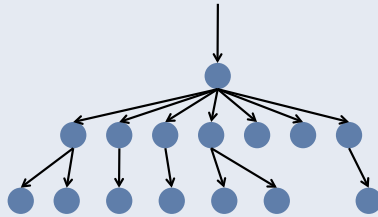
Eric Sun, Stanford University  
Itamar Rosenn, Facebook  
Cameron A. Marlow, Facebook  
Thomas M. Lento, Facebook

### Motivation

- How do ideas diffuse through a large social network?
- Prior models start with an isolated event and explore conditions under which this event triggers a global cascade
  - Little empirical evidence to assess the validity of the models
- We present the first empirical analysis of repeated large-scale diffusion over a global social network.

## Theory of the Influentials

- Popular Wisdom: it's all about the "influentials" (Malcolm Gladwell, etc.)
- Idea: reach a tiny group of influential people, and you'll reach everyone else through them for free



## Accidental Influencers

- Duncan Watts: anyone can be an "influencer".
- Ideas don't spread via influentials
  - Ideas spread like viruses: either you're susceptible, or you're not
- Success depends not on how persuasive the early adopter(s) are, but whether everyone else is easily persuaded



- Watts simulations: influential nodes are no more likely to trigger cascades than average nodes.

## Questions

- In most network models of diffusion, contagion is triggered by a fairly small number of sources: is this a good assumption?
- What are some characteristics of diffusion chains on Facebook?
- Can we use demographic or behavioral characteristics to predict the size of diffusion chains that a particular user will create?

## Data and Methodology

## Spreading Ideas on Facebook

The screenshot shows a Facebook News Feed with the following posts:

- Venky Iyer** is thinking we should measure statuses with george bush/barack obama in them over time. 42 minutes ago - Comment - Like
- Anjali Khurana** at 11:25pm January 16  
lexicon?  
Write a comment...
- Andreas Quandt** is having a nice beer after long day's work.  
about an hour ago - Comment - Like
- Thomas Lento** is pretty sure he can eat again. OM NOM NOM NOM!  
about an hour ago - Comment - Like
- Mary Ann Bailey** uploaded a mobile photo.  
Daisy is leaving me for emmie
- Wayne Chang** commented on **Mystery Slee's** photo.  
this was a triumph...
- Venky Iyer** posted a link.  
209.85.173.132  
Source: 209.85.173.132  
This is the html version of the file <http://www.pycon.it/static/pycon2/slides/containers.ppt>. Google automatically generates html versions of documents as we crawl the web.

## Empirical Study

- Focus on one type of action for an empirical analysis: Page fanning

The screenshot shows the Facebook page for the **Facebook Data Team**. The page includes the following elements:

- Header:** "data facebook" logo, "Facebook Data Team" name, and "Become a Fan" button.
- Navigation:** "Wall", "Info", and "Notes" tabs.
- Detailed Info:**
  - General Information:** "The Facebook Data team builds scalable platforms for the collection, management, and analysis of data."
    - "We use these platforms to help drive informed decisions in areas critical to the success of the company."
    - "We build tools and provide support for anyone at Facebook who would like to use our platforms to help make data-driven decisions or build data-intensive products and services."
- Fans:** "6 of 2,444 fans" and "See All" link.
- Profile Pictures:** Three small profile pictures of team members.

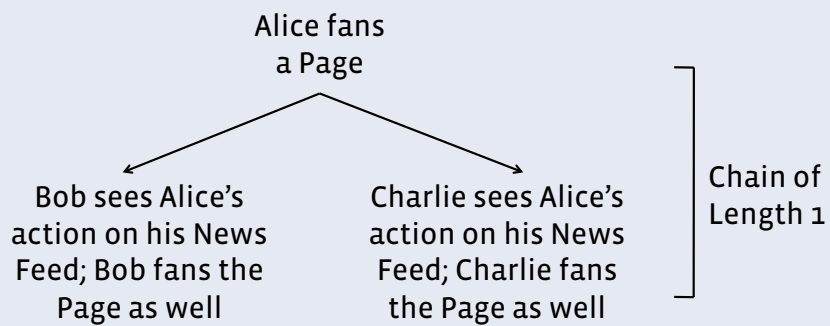
## Sample News Feed story



- Aside: current model of Page fanning...



## Mechanics of Page Fanning



## Data

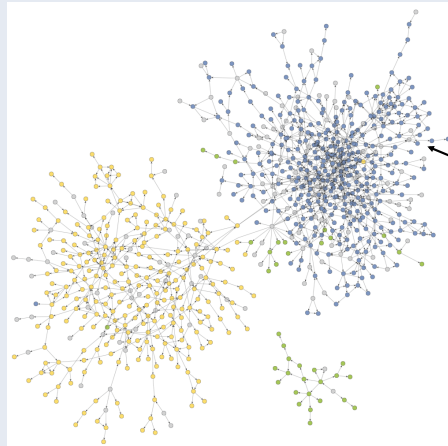
- Data include all actor → follower connections for 262,985 Facebook Pages between 2/19/08 and 8/19/08
- Main dataset: Page-level data
- Second dataset: select 10 random, representative Pages and analyze the users that start chains
  - Pages were at least 40 days old and had at least 5,000 fans

## Prediction Model

- Response : *max\_chain\_length*
- Predictors:
  - gender
  - log age
  - log Facebook age
  - log *feed\_exposure* (# friends who saw News Feed story)
  - log *friend\_count*
  - log *activity\_count* (wall posts + messages sent + photos added)
  - log popularity (controls for News Feed exposure)
- Method: zero-inflated negative binomial regression

# Results

## Large Connected Trees of Diffusion

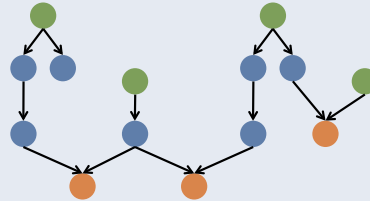


Link drawn if the follower fans a Page within 24 hours of first seeing a News Feed story that a friend has fanned the same Page

*Diffusion chain for Stripy, a cartoon popular in Bosnia (blue), Slovenia (yellow), and Croatia (green).*

## Large Clusters Not Started by “One Guy”

- Roughly 15% of fans in the biggest cluster of each Page are start points
  - The variability in this percentage becomes very small as # fans increases



- Clusters are formed when many short diffusion chains merge
- 86.4% of paths of Page diffusion involve at least four individuals
  - Compare to 38% in real-life study (Brown and Reingen 1987)

## Results of Chains Regression

Variable	<i>Intuition</i>	Fuddruckers	Cruise	Bolt	Zidane
(Intercept)	N/A	<b>2.574</b>	<b>1.419</b>	0.596	-0.105
age	?	<b>-0.537</b>	<b>-1.006</b>	<b>-0.386</b>	<b>-0.981</b>
gender==male	?	-0.011	-0.076	<b>-0.144</b>	<b>0.116</b>
Facebook_age	+	<b>-0.522</b>	-0.052	<b>-0.415</b>	<b>0.153</b>
activity_count	+	-0.102	<b>-0.142</b>	<b>-0.064</b>	<b>-0.100</b>
friend_count	+	<b>-0.220</b>	0.087	-0.023	0.009
feed_exposure	<i>control</i>	<b>1.279</b>	<b>1.008</b>	<b>0.860</b>	<b>1.053</b>
popularity	<i>control</i>	-0.014	<b>-0.245</b>	0.021	<b>-0.120</b>

- Recall: *max\_chain\_length* is the response variable
- Demographic characteristics not important
- Number of Facebook friends not important



## Conclusions

- Facebook News Feed enables long-lasting chains of diffusion that may reach many more people than real-life diffusion chains.
- The Facebook network is very connected: ideas with good receptiveness will attract wide, long connected clusters.
- Long chains are not a function of Facebook age, activity, users' demographics, or even # of friends: it's only related to exposure.

## Future Work

- Evaluate how accurately various theoretical models of diffusion account for the empirical phenomena uncovered
- Test experimental contagion events to better understand how different pieces of content and different start conditions determine the eventual structure of a diffusion cascade
- Test diffusion of other types of content

# Gesundheit!

## Modeling Contagion through Facebook News Feed

Eric Sun, Stanford University <esun@cs.stanford.edu>  
Itamar Rosenn, Facebook <itamar@facebook.com>  
Cameron A. Marlow, Facebook <cameron@facebook.com>  
Thomas M. Lento, Facebook <lento@facebook.com>