

**USC Viterbi**  
School of Engineering

# Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks

**Kristina Lerman**

**Rumi Ghosh**

USC Information Sciences Institute



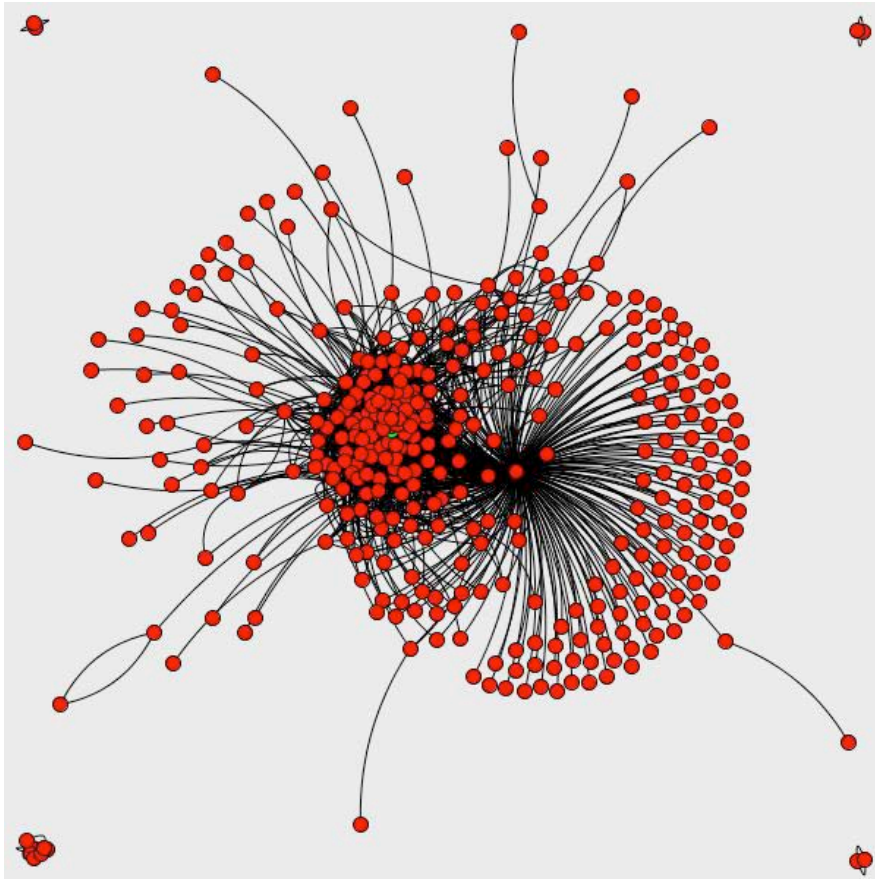
# Online Social Networks



Online social networks have become important channels for the spread of timely and relevant information

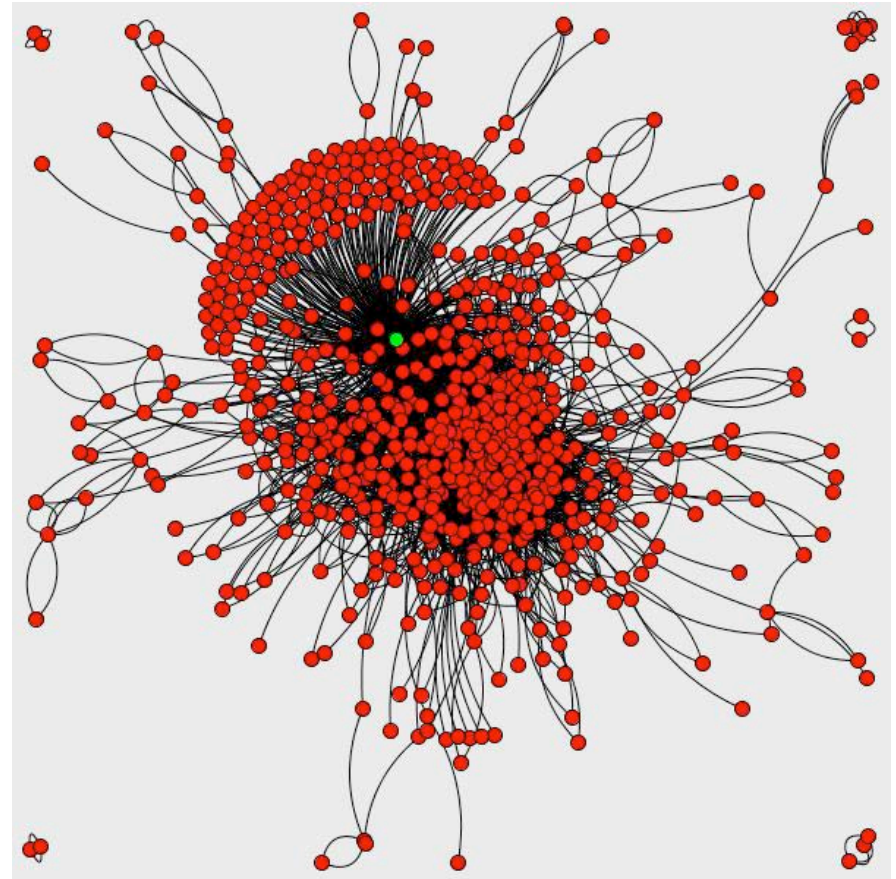
# Information flow on networks

Network 1



<http://www.flickr.com/photos/27318782@N03/4639307471/>

Network 2



[http://www.flickr.com/photos/27318782@N03/4639307923](http://www.flickr.com/photos/27318782@N03/4639307923/)

## Dynamics of Social Information

### How does information spread on online social networks?

- How far and how fast does information flow on networks?
- What factors influence its spread?
- How does the network structure affect dynamics of information flow?
- What does this tell us about the quality of information?

### Study these questions through a comparative empirical analysis of two social news networks

- Using URLs as markers for tracking the flow of information



## Social news: Digg

Users submit and vote for (digg) news stories

Users join networks to see

- Stories friends submit
- Stories friends vote for

Digg features stories with most votes on its front page

The screenshot shows the Digg website interface. At the top, there is a navigation bar with the Digg logo and links for Profile, Friends' Activity, and Submit New. Below this is a category bar with Technology, World & Business, Science, Gaming, Lifestyle, and Entertainment. A secondary bar shows Popular and Upcoming sections, with sub-links for News, Videos, Images, and a Customize option.

The main content area displays a list of news stories, each with a digg count, a title, a source, a brief description, and interaction options (Comments, Share, Bury, and a 'made popular' button).

- 319 diggs**: [FDA to be aggressive in tobacco regulation](#) (courier-journal.com) — If there is any doubt about how aggressive the federal Food and Drug Administration intends to be in regulating tobacco, take a look at a letter the agency sent out last week. **173 Comments** | **badqat made popular 14hr 35 min ago**
- 408 diggs**: [Can a Daily Pill Really Boost Your Brain Power?](#) (guardian.co.uk) — In America, university students are taking illegally obtained prescription drugs to make them more intelligent. Here, an investigation into the brave new world of neuro enhancement... **83 Comments** | **openthink made popular 14hr 45 min ago**
- 2215 diggs**: [What happens when your Mom cancels your WoW account...](#) (revision3.com) - Parents just don't understand. We don't either - what WAS he trying to do with that remote control? **Share** | **Bury**
- 168 diggs**: [Three suspects arrested in U.S. terrorism probe](#) (reuters.com) — A Colorado man, his father and an accused accomplice in New York were arrested on Saturday and charged with lying to federal agents about a plot to blow up unspecified targets in the United States, the U.S. Department of Justice said. **60 Comments** | **11 cuisinart made popular 14hr 54min ago**

## Social news: Twitter + Tweetmeme

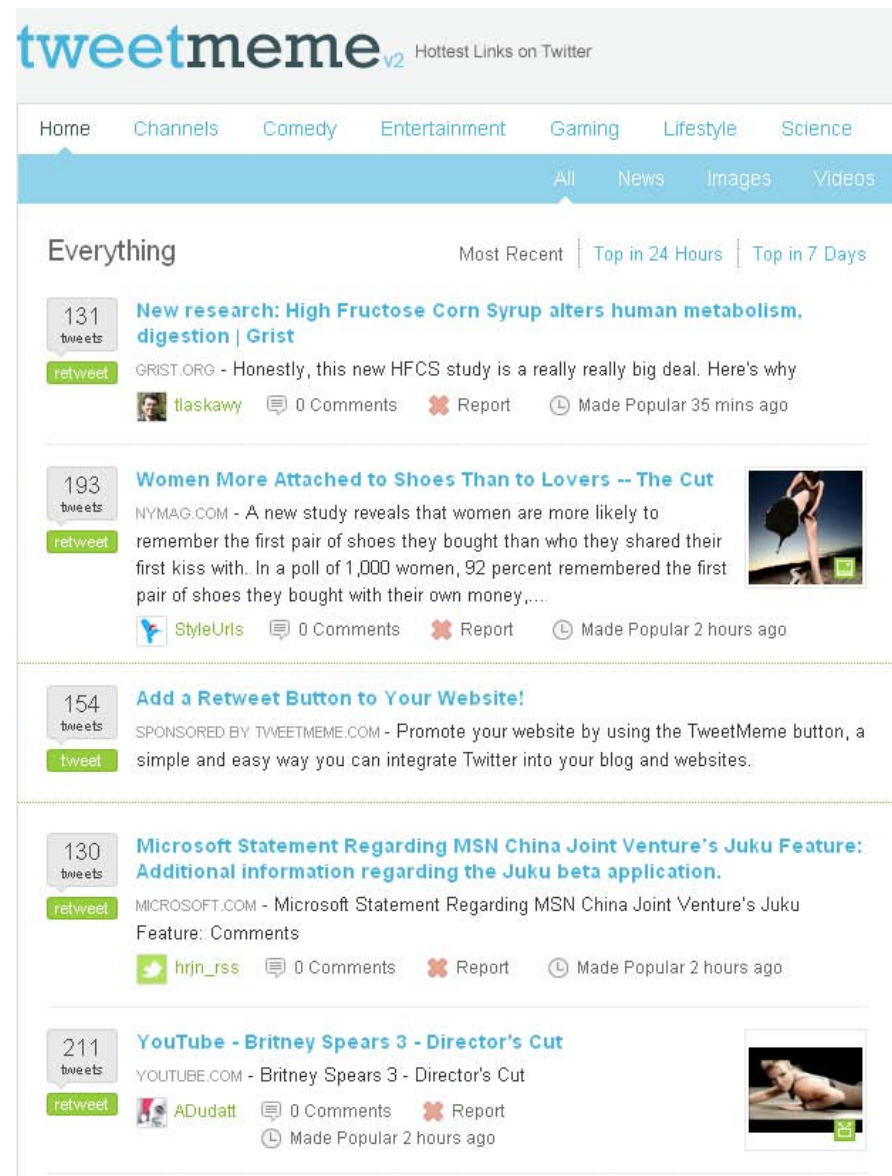
Users tweet and retweet\* URLs to news stories

\*‘Retweet’ = tweet someone else’s post  
 “RT @x failed bomb plot <http://bit.ly/xmas09>”

Users join networks to see

- Tweets by users they follow
- Retweets by users they follow

Tweetmeme aggregates all tweets and features most retweeted URLs on its front page



## Social news: Data sets

### Digg

- **Stories**
  - 3,553 stories, promoted in June, 2009
    - Time submitted, promoted
  - Votes for each story
    - Time of the vote
    - Name of voter
- **Active users**
  - 139,409 who voted for at least one story
  - 71,834 of them following at least one user
  - 258,220 links
  - → fan network

### Twitter

- **Stories**
  - 398 most retweeted stories 6/11/09—7/3/09
    - extracted from Tweetmeme
  - Retweets of each story
    - <1000 most recent retweets
    - Time of retweet & user name
- **Active users**
  - 137,582 who retweeted at least one story
  - Following/follower relations through Twitter API
  - 6,200,051 links
  - → follower network

## Questions

### Usability of Social News

- Do people use Digg and Twitter the same way?
- What effect to differences in the user interface have?

### Dynamics of Social News

- How far does information spread on networks?
- How fast does information spread on networks?
- What is the role of network structure?



## Basic terms

### Submitter

- user who submitted link to a story
- user who first tweeted link to a story

### Vote

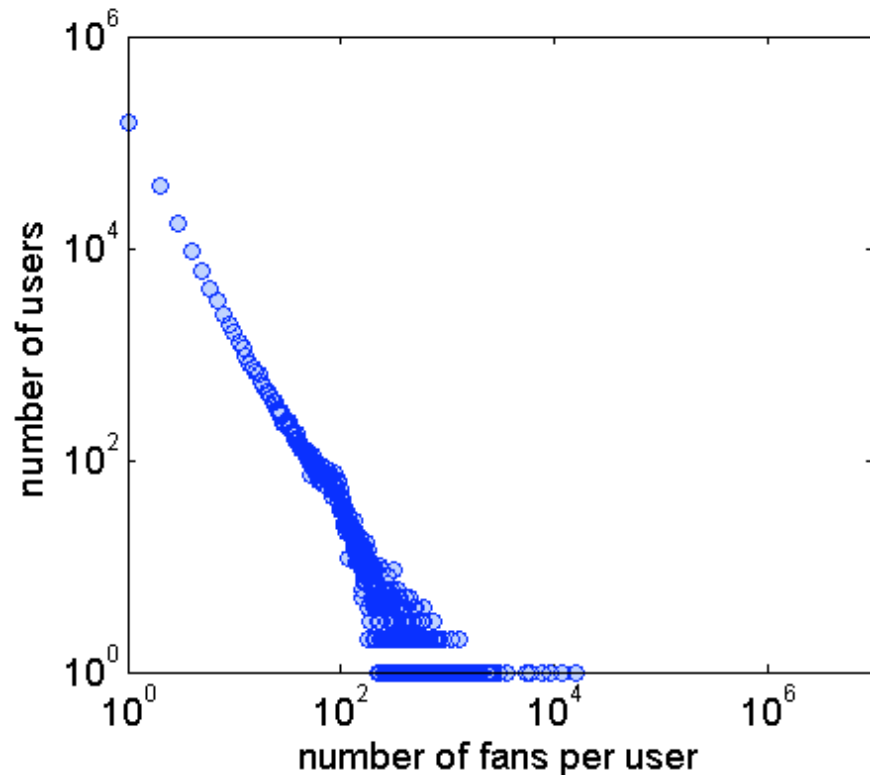
- digg
- retweet

### Fan of user A

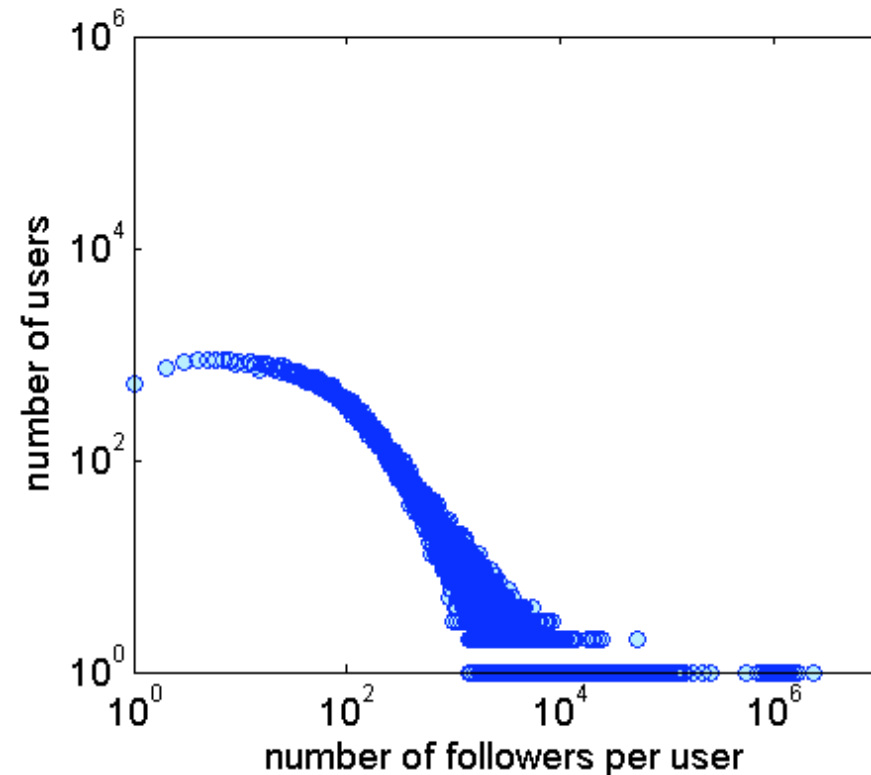
- user watching A's activity on Digg
- user following A on Twitter

## User activity: distribution of fans

### Digg



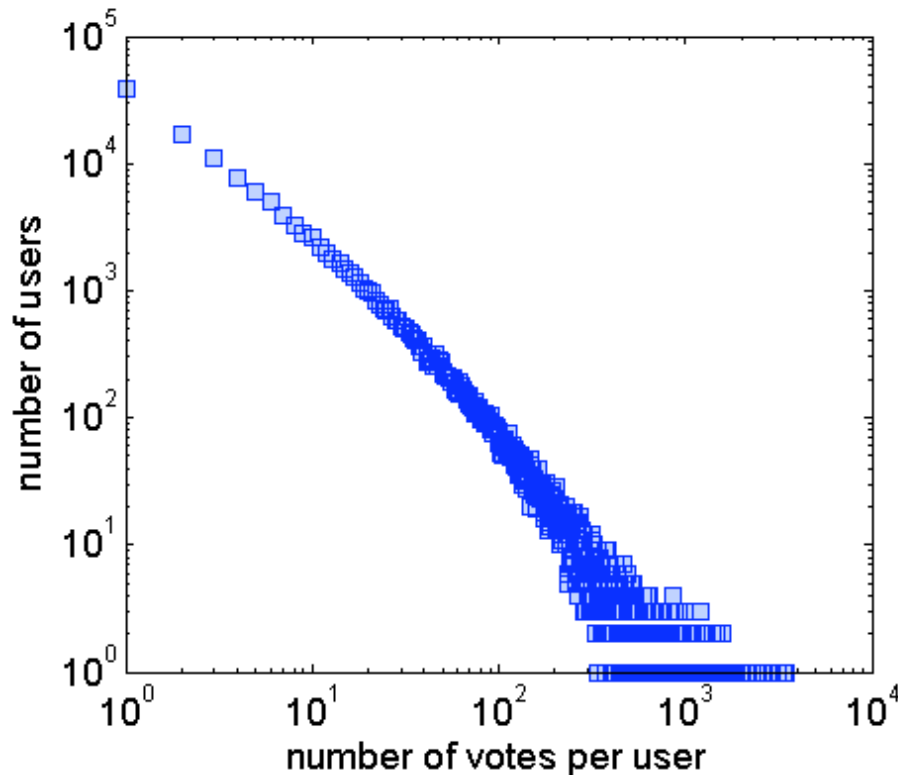
### Twitter



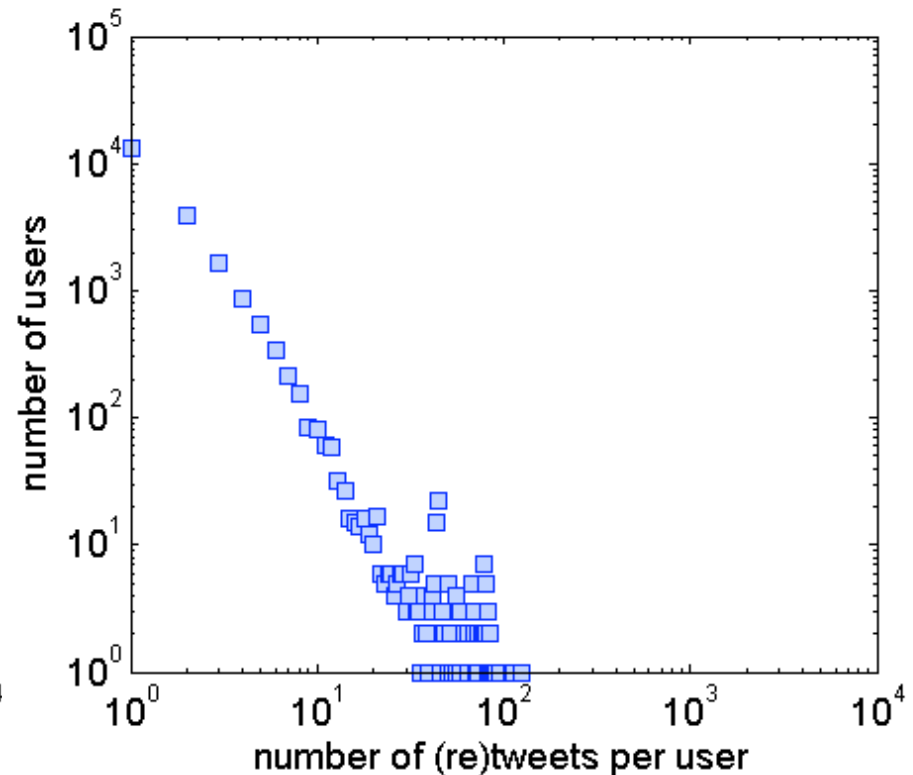
- Typical number of followers on Twitter  $\sim 10$ , but can be millions
- No typical number of fans on Digg – “long tail”

## User activity: distribution of voting

### Digg



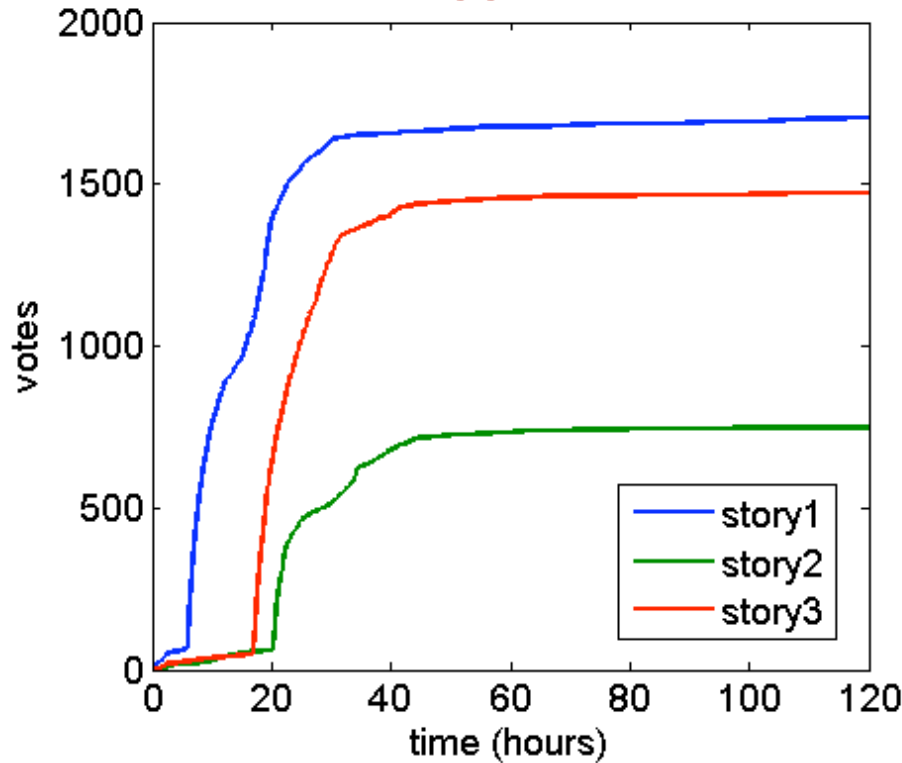
### Twitter



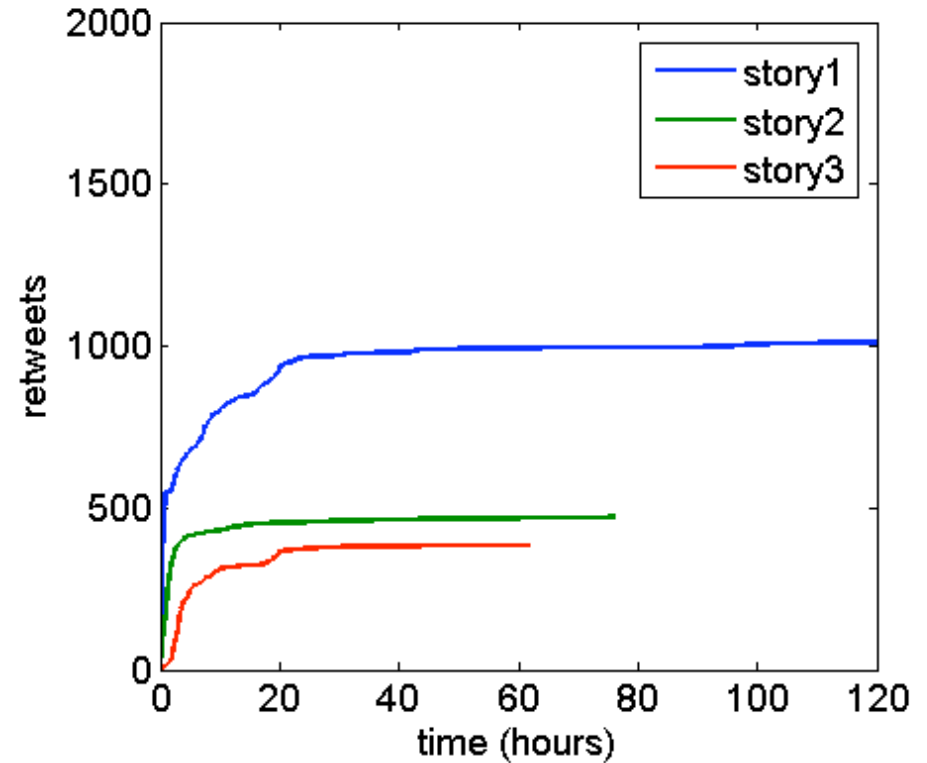
- “Long tail” distribution of user activity
- Difference in slope related to effort of activity [cf Wilkinson 2008]

# Dynamics of stories

Digg



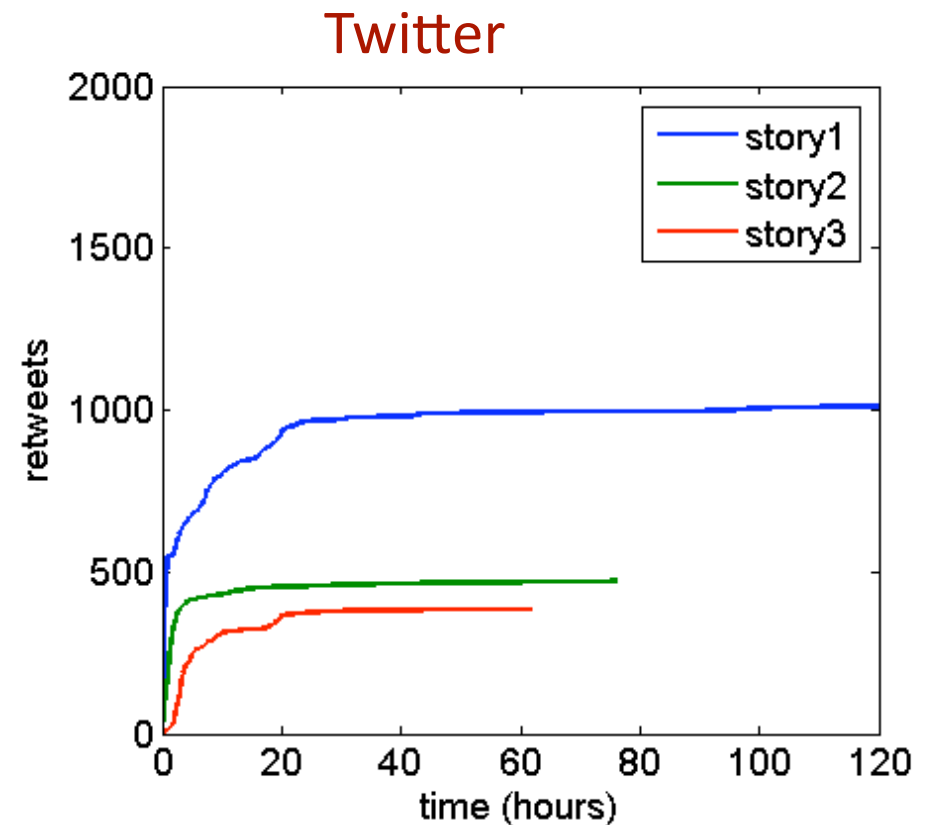
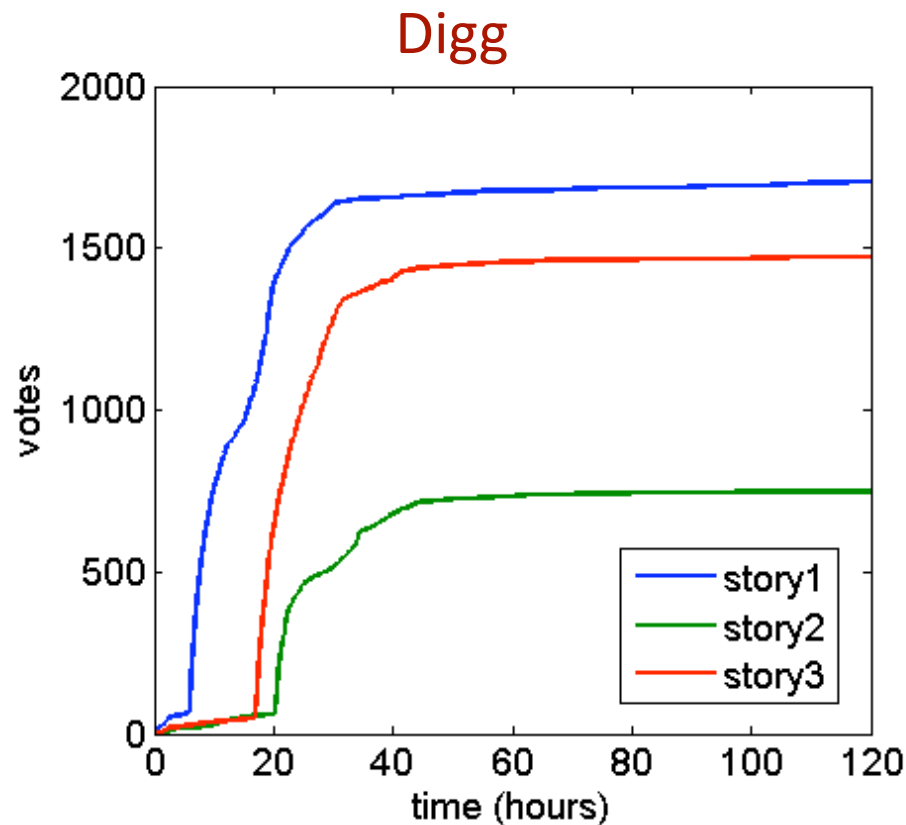
Twitter



- 1: U.S. Government Asks Twitter to Stay Up for #IranElection
- 2: Western Corporations Helped Censor Iranian Internet
- 3: Iranian clerics defy ayatollah, join protests

- 1: US gov asks twitter to stay up
- 2: Iran Has Built a Censorship Monster with help of west tech
- 3: Clerics join Iran's anti-government protests - CNN.com

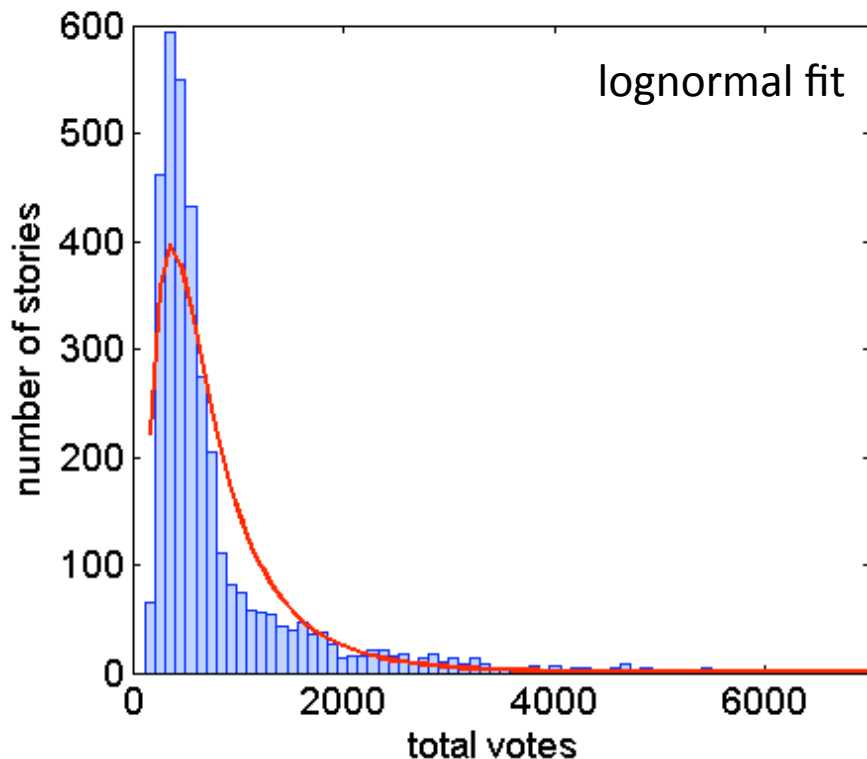
## Dynamics of stories



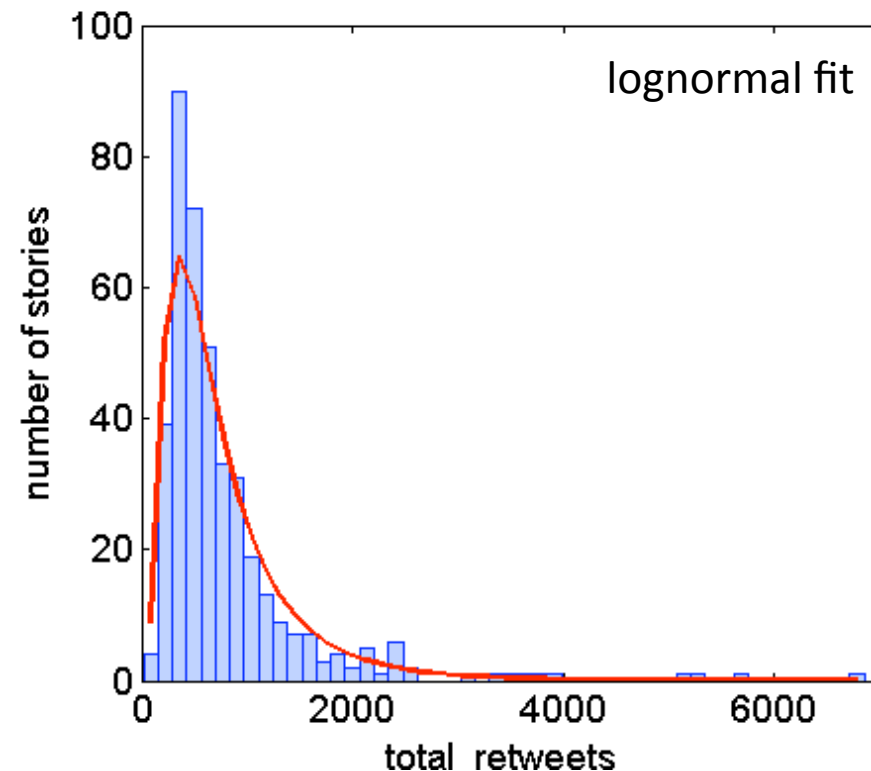
- Two distinct phases for Digg stories: upcoming and promoted
- Number of votes on both sites saturates after one day
- Saturated value reflects story popularity

# Popularity of stories

## Digg

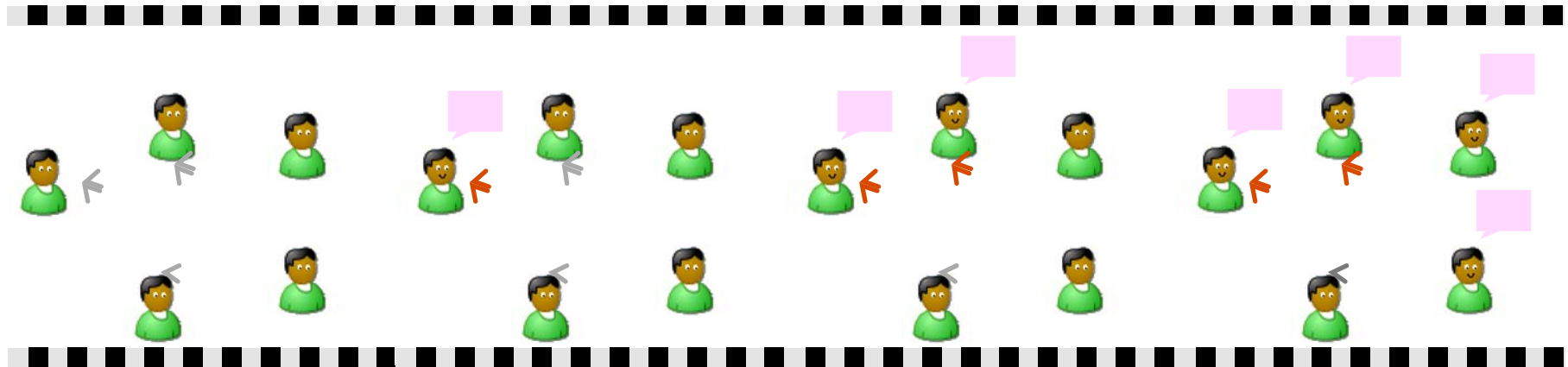


## Twitter



- Aggregate over all stories to factor out influence of submitter and story quality
- “Inequality of popularity” – some stories much more popular than others  
*cf* social influence study of [Salganik, Dodds & Watts, 2006]

# Information flow on networks

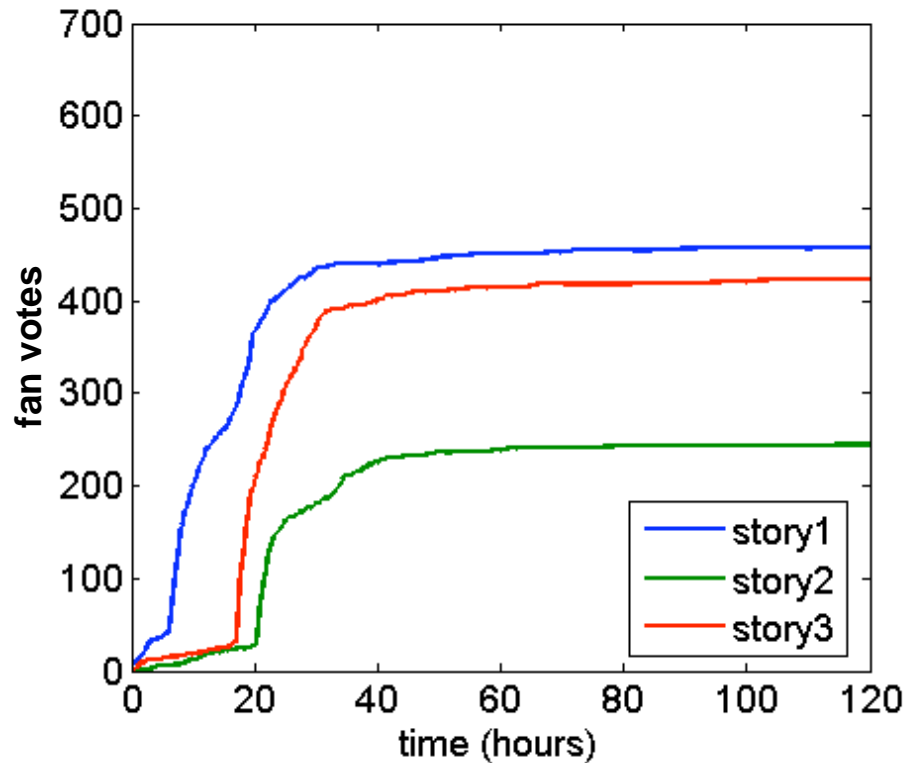


Information spreads on a network as fans (followers) vote for (retweet) stories their friends submit or vote for.

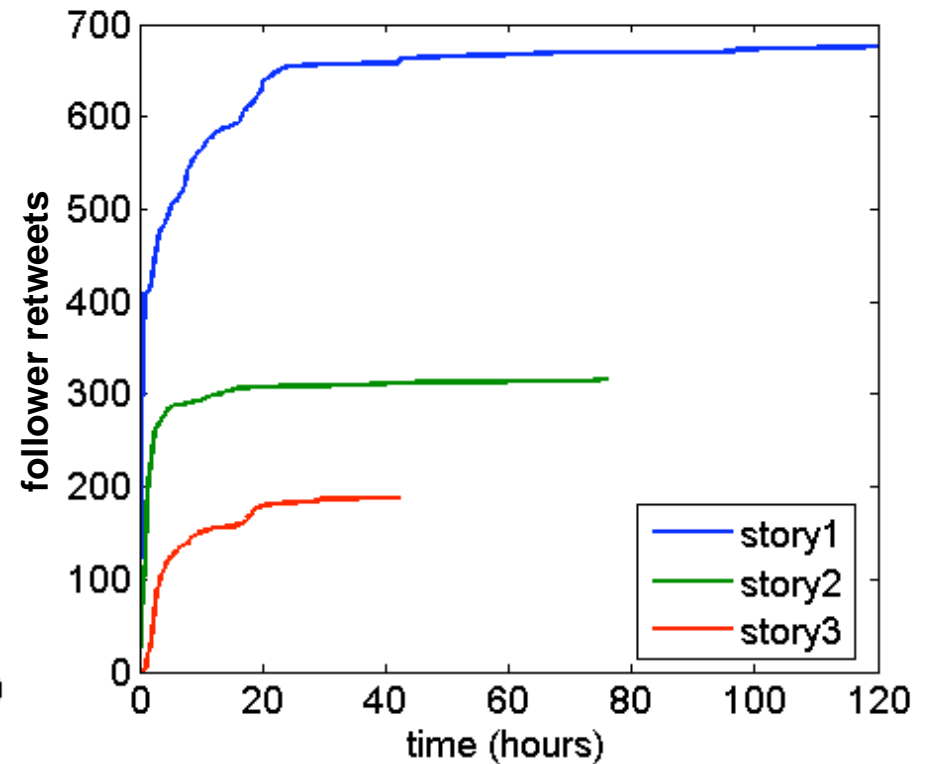
... **fan votes** – i.e., votes from fans

# Dynamics of information spread on networks

## Digg



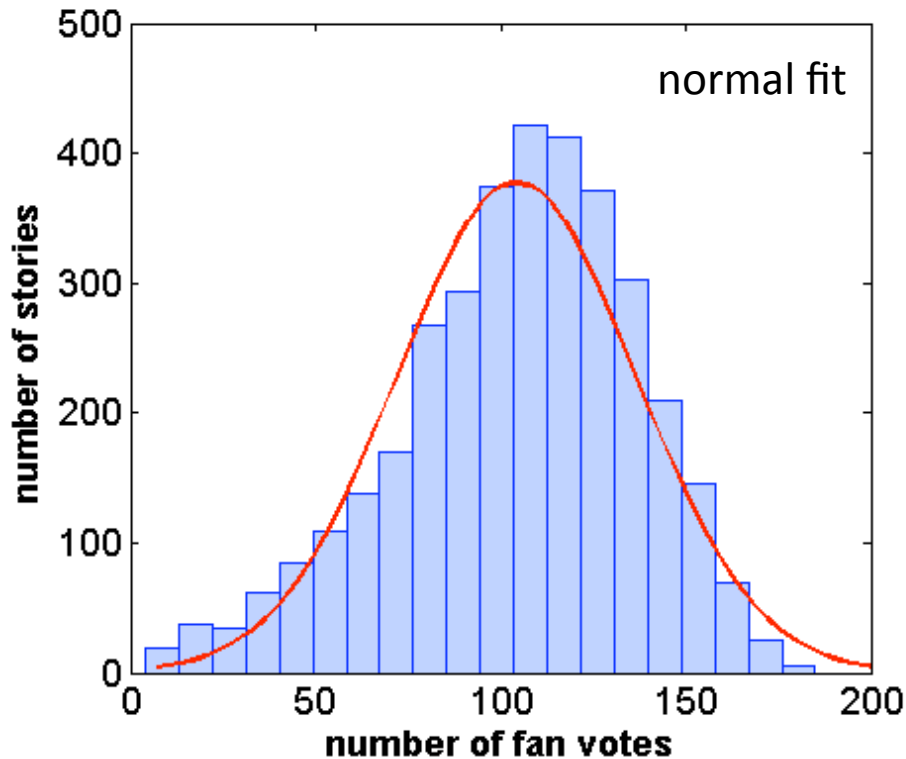
## Twitter



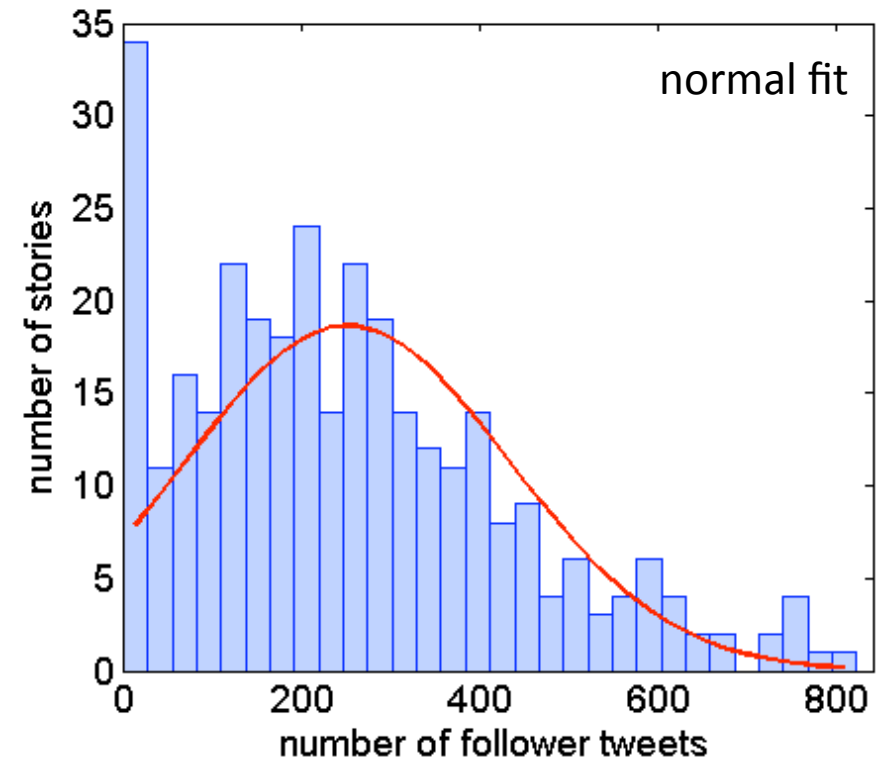
- Evolution of fan votes qualitatively similar to that of all votes

## How far does information spread on networks?

### Digg



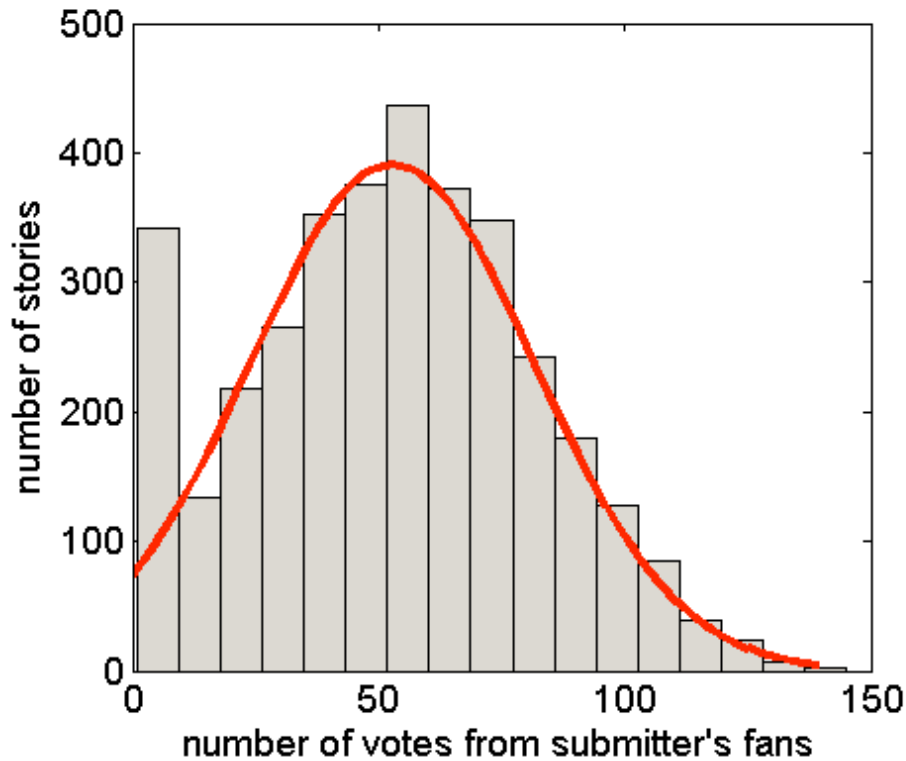
### Twitter



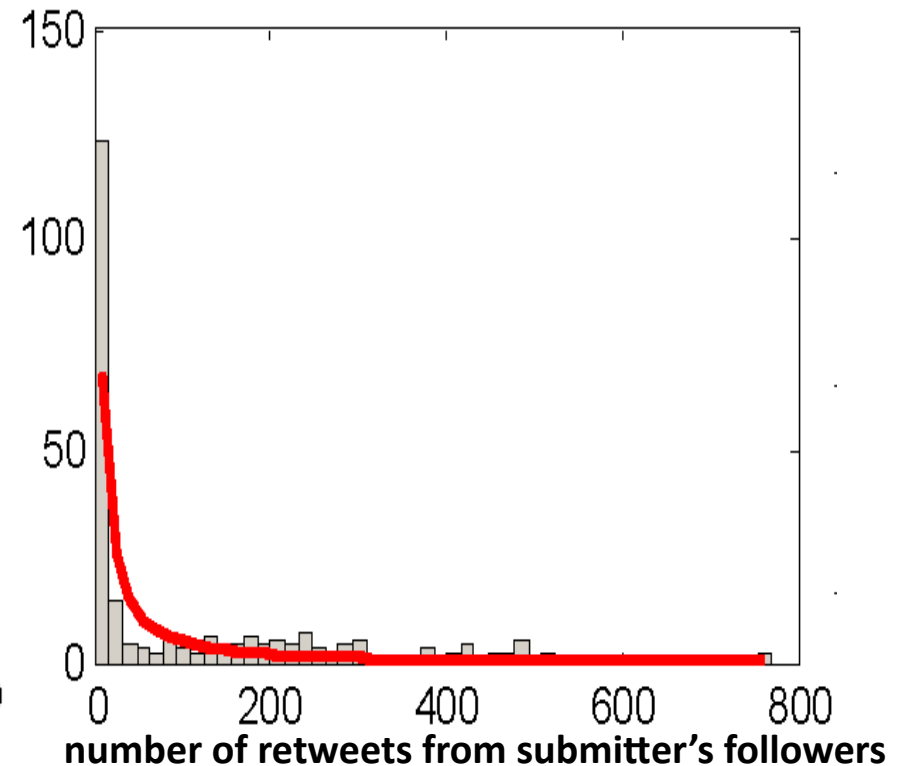
- “Inequality of popularity” no longer observed (social influence accounted for?)
- News spreads farther on Twitter social networks
- On Digg, all stories receive fan votes

## How far does information spread among submitter's fans?

### Digg

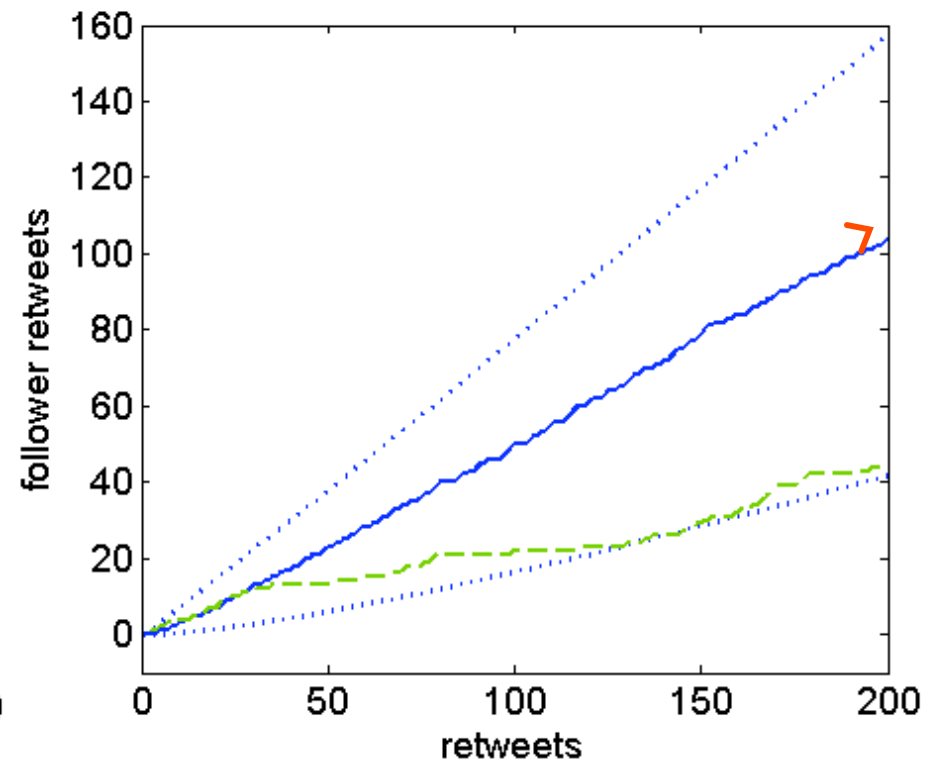
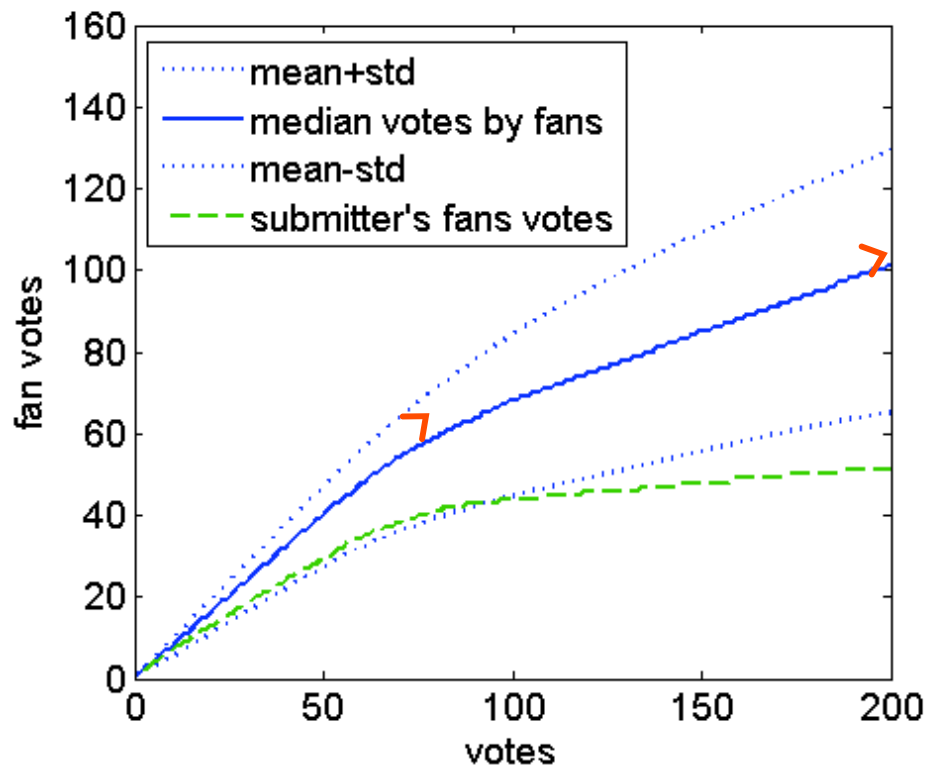


### Twitter



- Most stories on Twitter are never retweeted by submitter's followers

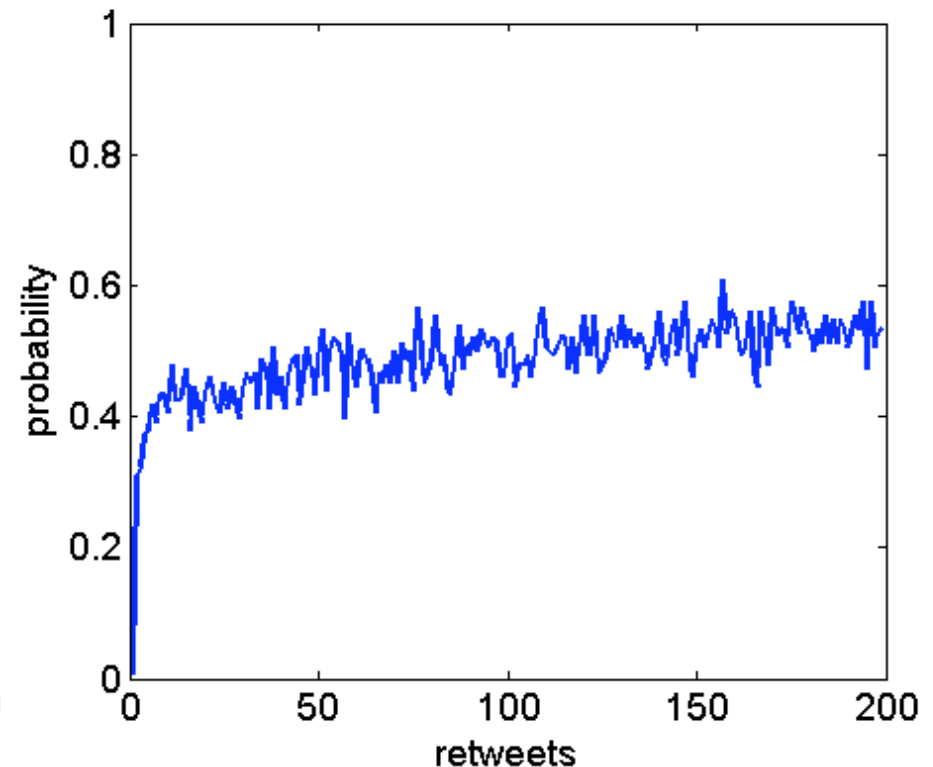
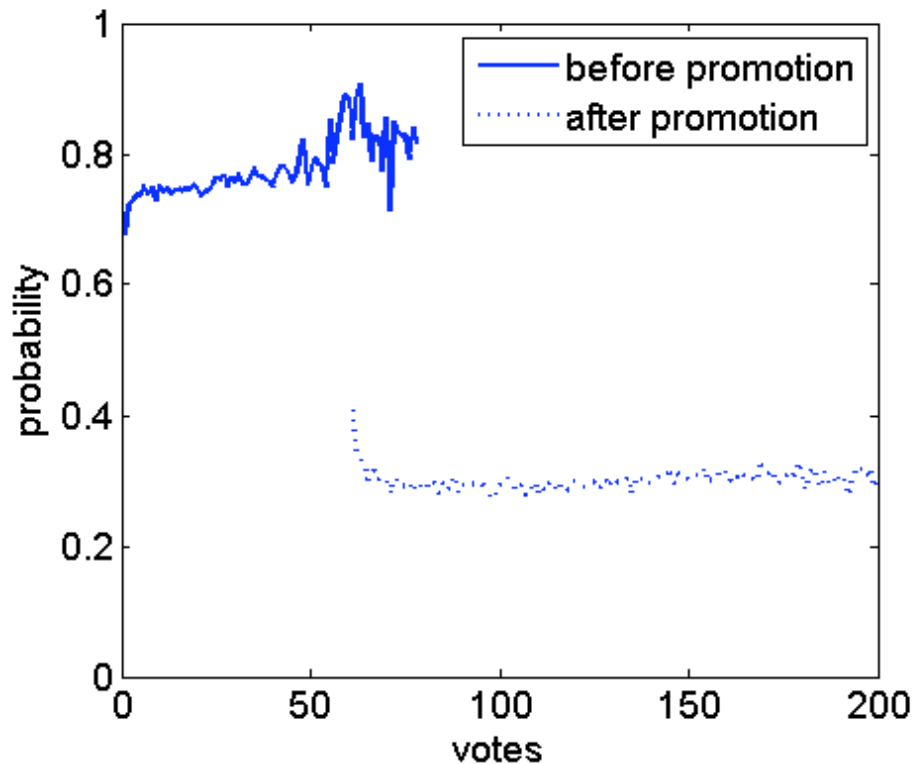
## How fast does information spread on networks?



- Two distinct phases on Digg: stories spread faster through the network before promotion than afterwards
- Twitter stories spread at a uniform rate, but with greater variability

# How fast does information spread on networks?

Probability next vote is from a fan



- Two phases on Digg: stories spread faster before promotion than after
- Twitter stories spread slower than Digg stories before promotion, but faster than promoted stories

## Network structure

### Network sample size

- Digg: 279,725 users
- Twitter: 6,200,051 users

### Network density

- Fraction of reciprocal links (i.e., mutual fans)
  - Digg:  $f_m = 3.20 \times 10^{-6}$
  - Twitter:  $f_m = 2.07 \times 10^{-7}$
- Modified clustering coefficient (fraction of cycles, e.g.,  $A \rightarrow B \rightarrow C \rightarrow A$ )
  - Digg:  $f_c = 7.60 \times 10^{-12}$
  - Twitter:  $f_c = 1.92 \times 10^{-14}$

Digg network is denser, more inter-connected than Twitter's

## Summary of results

### Network structure and information flow

- Digg's network is denser than Twitter's
  - News spreads faster initially through Digg's network than Twitter's
  - But, it does not spread as far as on Twitter
  - Fans who vote for a story on Digg are also submitter's fans
- Twitter's network is sparse
  - Fans unconnected to submitter help spread the story

### User interface and information flow

- Before promotion, Digg stories spread mainly through the network
  - Similar to story spread on Twitter
- After promotion, stories spread mainly outside the network
  - Promotion increases story visibility outside the network
  - No equivalent mechanism on Twitter

## Related work

- **Information flow on social networks**
  - Email chains [Wu, Huberman, et al, 2004]
    - Email forwarding chains terminate after a few steps
    - Information flow is slowed by decay in similarity among individuals in a social network
  - Product recommendations [Leskovec, Adamic & Huberman, 2006]
    - Word-of-mouth recommendation chains terminate after 1 or 2 steps

We find significantly large information spreading chains
- **Information diffusion through blogosphere**
  - Spread of topics through blogosphere [Gruhl & Liben-Nowell 2004]
  - Information cascades through blog posts [Leskovec et al, 2007]
    - Power law distribution of cascade sizes

We find a normal distribution of cascade sizes (number of fan votes)
  - **Networks derived from observed interactions, rather than declared links (as on Digg, Twitter)**

## Conclusion

### Comparative empirical analysis of online social networks and social news on Digg and Twitter

- Similarities
  - Networks are used on both sites to spread news
  - Information flow on Digg before promotion is similar to information flow on Twitter
- Differences in dynamics can be explained by user interface and network density
  - Digg's user interface gives high visibility to promoted stories
    - Promotion slows information spread
  - Digg's network is denser than Twitter's
    - Affects how fast and how far news spreads
- Dynamics of information flow as a gauge of information quality
  - Social dynamics of Digg [Hogg & Lerman, poster][Lerman & Hogg, WWW 2010]