Malicious Behavior on the Web: Characterization and Detection

Srijan Kumar (@srijankr)
Justin Cheng (@jcccf)
Jure Leskovec (@jure)

Slides are available at http://snap.stanford.edu/www2017tutorial/

Social media has brought us closer together…

Vieweg, et al. (2010); Kittur, et al. (2013); Burke & Kraut (2016)

Social Media Allows Us..

- Social Media disseminates through Social interaction
- Web is no longer a static library that people passively browse
- Web is a place where people:
  - Consume and create content
  - Interact with other people:
    - Internet forums, Blogs, Social networks, Twitter, Wikis, Podcasts, Slide sharing, Bookmark sharing, Product reviews, Comments, …

Social Media: Opportunities

- Any user can share and contribute content, express opinions, link to others

- <u>This means:</u> Can data-mine opinions and behaviors of millions of users to gain insights into:
  - Human behavior
  - Marketing analytics
  - Product sentiment

Real-time citizen journalism

- **Citizen journalism** provides more valuable information than newswire services
- Challenge:
  - Many redundant posts, users have to wade through hundreds of posts to locate useful information
- Goal:
  - Mine this data in real-time and produce well organized summaries

Applications: Reputation management

- **Consumer Brand Analytics**
  - What are people saying about our brand?
- **Marketing Communications**
  - Significant spending on marketing, advertising: Companies trying to position their products
  - Brand analytics helps to determine whether such campaigns are effective
- **Product reviews**
- Automatically mine product reviews for information on product features, new requests, …
  - Easy to use, Light weight, Sturdy, Good price, …

…but there's also a dark side to social media

"On the Internet, nobody knows you're a dog."

Not everyone has good intentions...

How Trolls Are Ruining the Internet

When Will the Internet Be Safe for Women?

FAKE NEWS IS ABOUT TO GET EVEN SCARIER THAN YOU EVER DREAMED

Time (2016); The Atlantic (2016); Vanity Fair (2017)

Antisocial behavior is prevalent online

40%

60%

■ Harassed
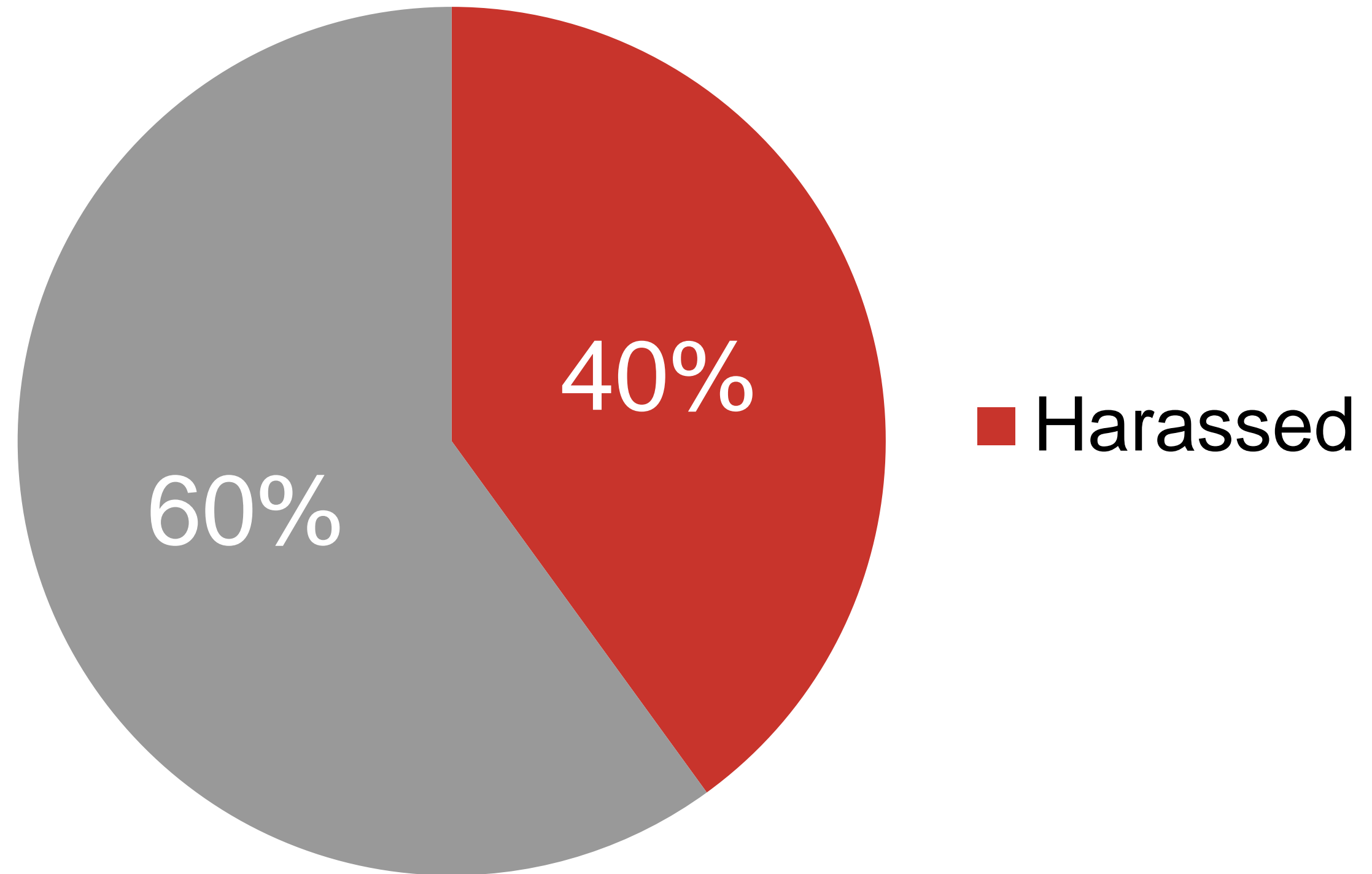
Pew Research (2014)

Antisocial behavior is widespread

39% of experts said they expect the online future will be "more shaped" by negative activities

Pew Research (2017)

Not everything is true...

Growing Number of Social Media Users are Misinformed About Current Events

- People who rely mainly on social media for news are often misinformed on current events or have formed an opinion on a topic without researching or fact checking.

- Journalists sometimes report actual news stories that are based on inaccurate social media updates.

More social media users rely on links from friends/family contacts

Facebook news driven by friends and family

Percent who get most of their news links from...

	Friends/family	News orgs/ journalists
FACEBOOK	70%	13
TWITTER	36%	27
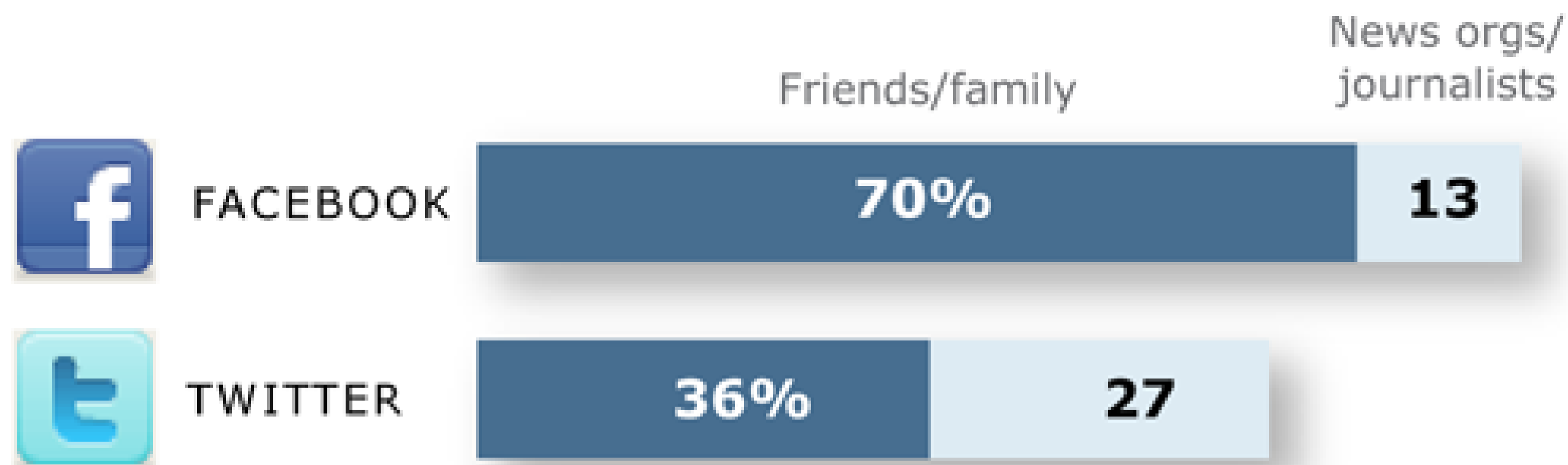
Graphic Source: Pew Research Center's Project for Excellence in Journalism / 2012 State of the News Media

While social media can help break real news...

- In 2009, Twitter user Janis Krums (@jkrums) tweeted photos of the US Airways "Miracle on the Hudson" plane that crashed in New York's Hudson River. He was on one of the ferries that responded to the crash site before traditional news outlets knew about the incident.

There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy.

880,678   1486 days ago

- SOURCE: Social Media Today
http://socialmediatoday.com/dlawrence/475239/social-media-now-major-source-breaking-news

While social media can help break real news...
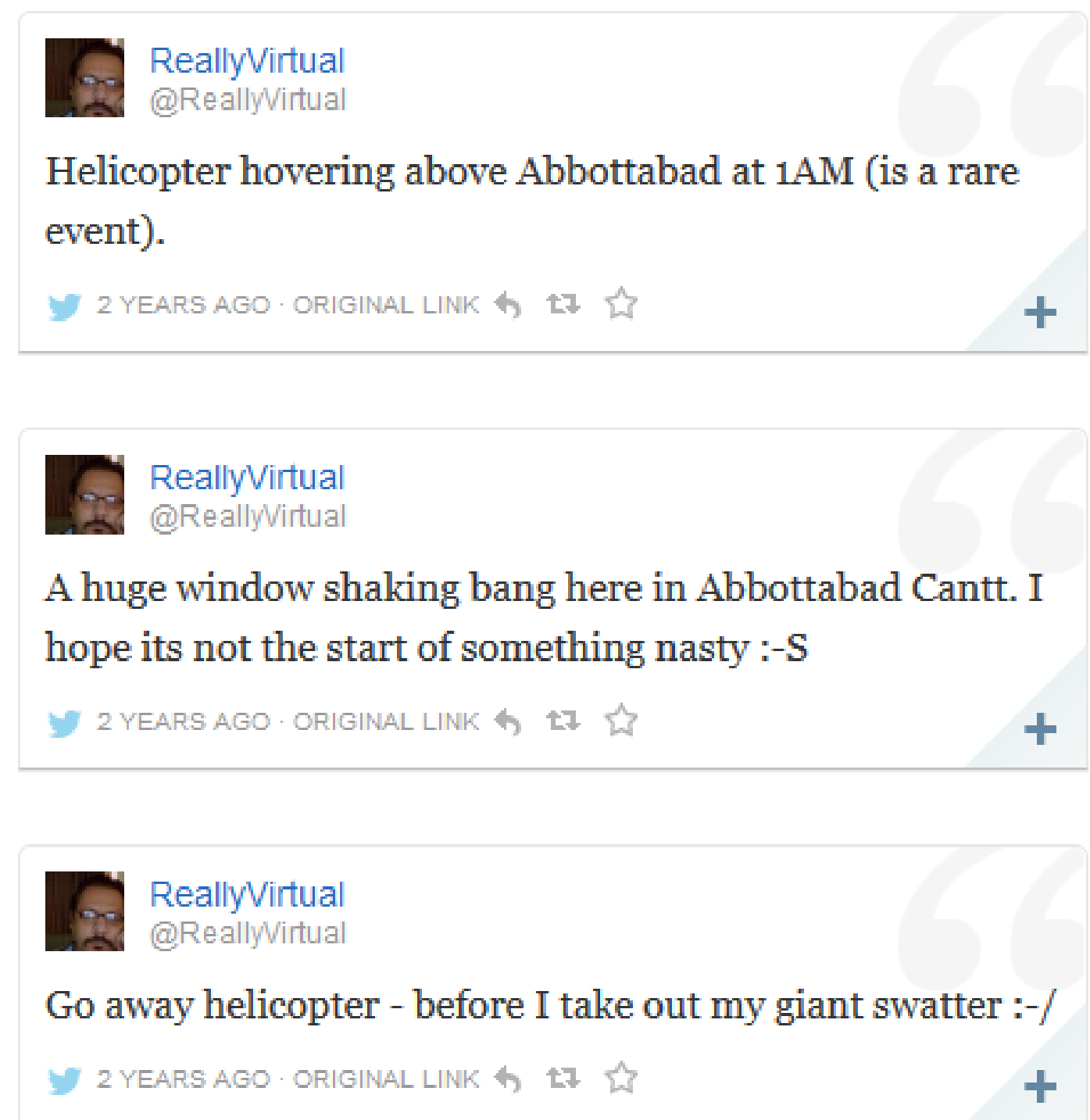
- In 2011, Twitter user @reallyvirtual tweeted about a helicopter annoying him over his home in Abbottabad, Pakistan, at 1am local time. He was live-tweeting about the US raid that killed Osama Bin Laden.

  - SOURCE: Washington Post.
    http://www.washingtonpost.com/blogs/blogpost/post/sohaib-athar-tweeted-the-attack-on-osama-bin-laden%20--%20without-knowing-it/2011/05/02/AF4c9xXF_blog.html

ReallyVirtual
@ReallyVirtual

Helicopter hovering above Abbottabad at 1AM (is a rare event).

2 YEARS AGO · ORIGINAL LINK

ReallyVirtual
@ReallyVirtual

A huge window shaking bang here in Abbottabad Cantt. I hope its not the start of something nasty :-S

2 YEARS AGO · ORIGINAL LINK

ReallyVirtual
@ReallyVirtual

Go away helicopter - before I take out my giant swatter :-/

2 YEARS AGO · ORIGINAL LINK

…Not everything published online is necessarily true.

- Jackie Chan died in March 2011…according to numerous Facebook posts.

  The actor then had to release a statement, via his Facebook page, to prove that the rumors were false and that he was alive.

成龍 Jackie Chan
March 29, 2011

Jackie is alive and well. He did not suffer a heart attack and die, as was reported on many social networking sites and in online news reports.

Jackie is fine and is busy preparing for the filming of his next movie.

Like · Comment · Share   60,054   8,905   1,840

…Not everything published online is necessarily true.

■Morgan Freeman died in December 2010 and the news was credited to CNN… according to numerous "tweets" on Twitter.

CNN then had to clarify that they never reported the news and that Morgan Freeman was, in fact, alive.

Source: Mediate
http://www.mediaite.com/online/cnn-says-no-it-never-sent-out-tweet-reporting-death-of-actor-morgan-freeman-whos-alive/

CNN did not report Morgan Freeman death. Rumor is false. CNN will aggressively investigate this hoax.

3 minutes ago via web
Retweeted by 69 people

Reply   Retweet

CNN   CNN

Fake News Goes "Viral" Online

- Many social media users "retweet", "share", and "like" these erroneous reports.

- These users include average citizens and journalists who don't fact-check before spreading the news.
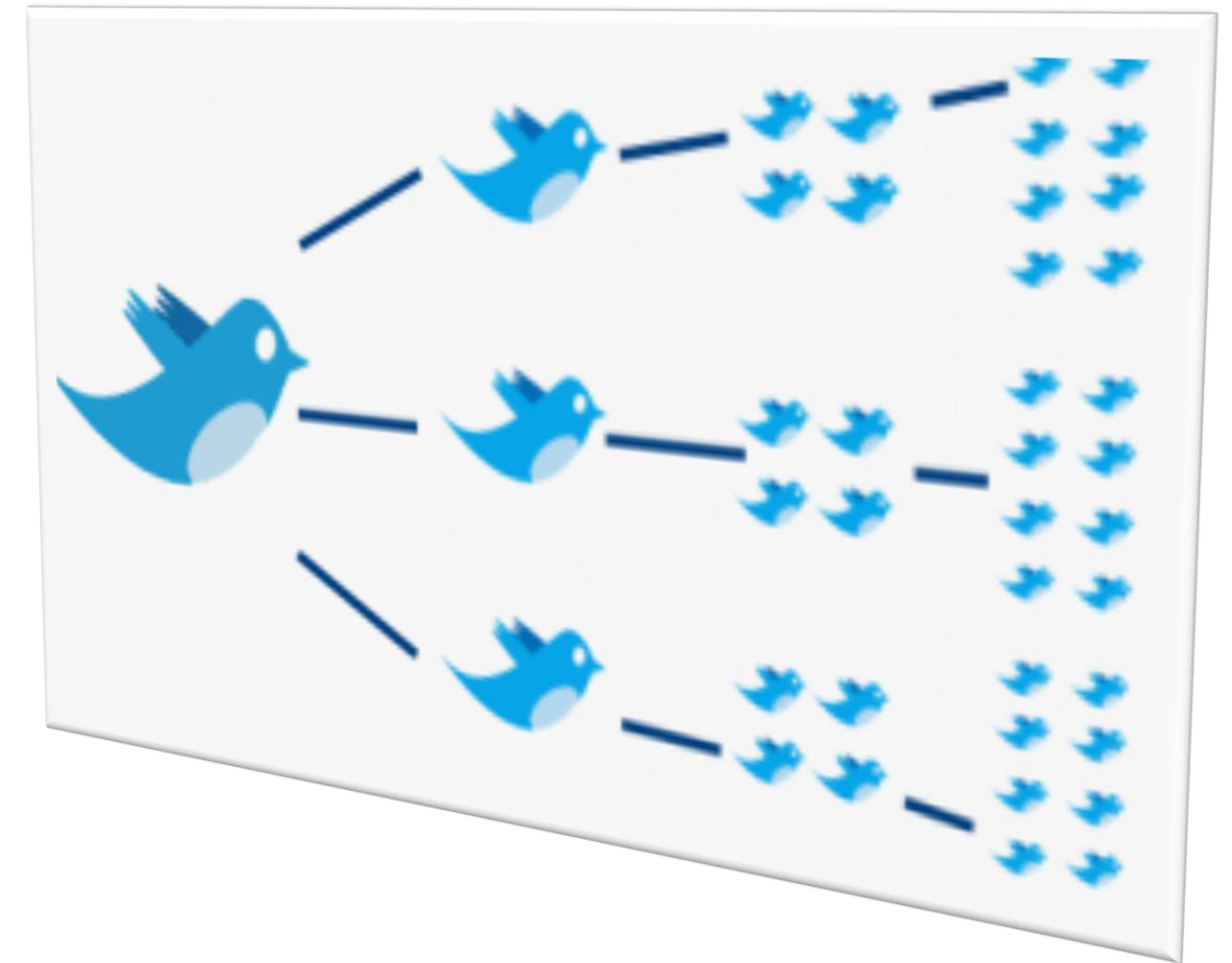
This tutorial:
Antisocial Behavior on the Web: Characterization and Detection

This tutorial: Goals

Towards design of better social systems

Can we use these insights to design better online interactions?

- How do we identify trolls/vandals/rumors and/or predict antisocial behavior?
  - What kinds of interventions could minimize antisocial behavior?

Tutorial Outline

Malicious users

Trolling Sockpuppets Vandals

Misinformation

Fake reviews Hoaxes

http://snap.stanford.edu/www2017tutorial

Goals of this Tutorial

- Discuss a set of methods for analyzing malicious behavior online

- Summarize insights obtained by various studies into different aspects of malicious behavior online:
  - Malicious users
  - Malicious content
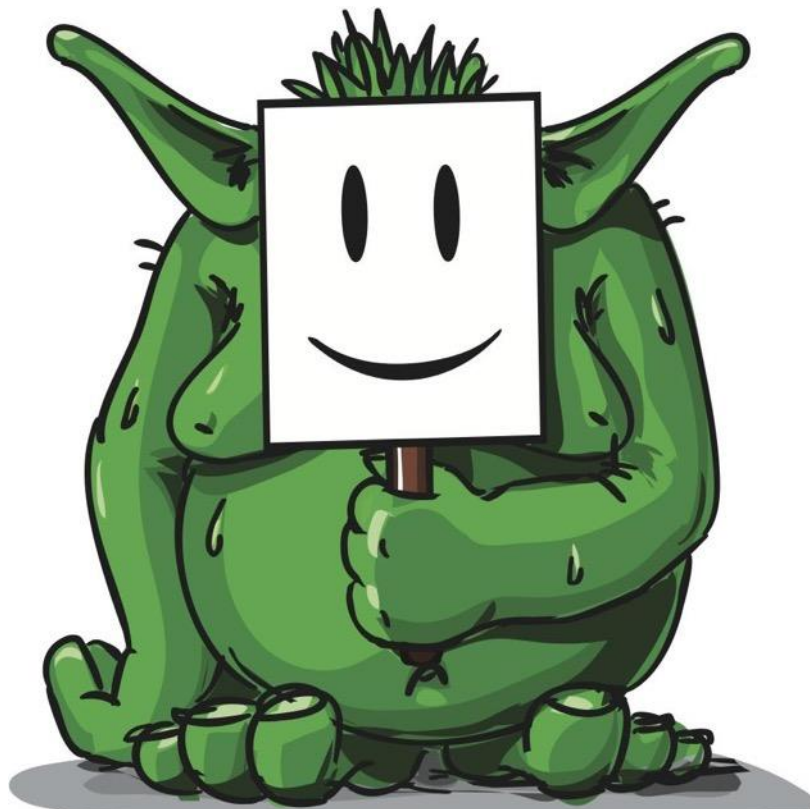
Challenges in analyzing antisocial behavior

Population imbalance

Smaller propotion of behavior (< 10%) is antisocial
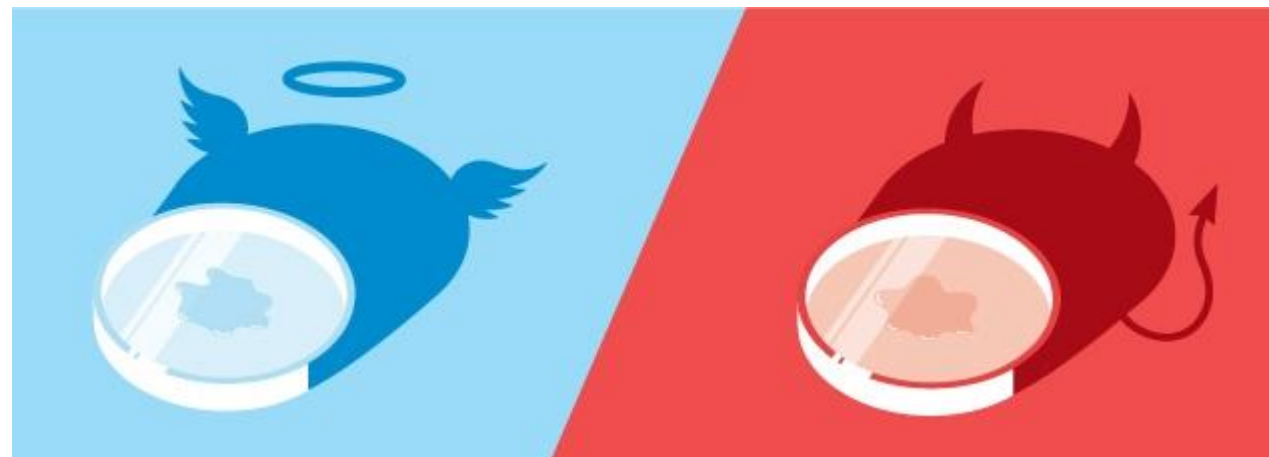
Limited labels

Little known information about antisocial behavior

Deceptive behavior

Antisocial behavior tends to masquerade as benign