



KDD 2017

Halifax, Nova Scotia - Canada
August 13 - 17, 2017

Tutorial: Data-Driven Approaches towards Malicious Behavior Modeling



Meng Jiang
University of Notre Dame



Srijan Kumar
Stanford University



Christos Faloutsos
Carnegie Mellon University



V.S. Subrahmanian
University of Maryland, College Park

Tutorial link: <http://bit.ly/kdd2017>

Web: A platform for everyone





Web and social media have brought
us closer together...



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

Web allows...

- Web enables social interaction
- Web is no longer a static library that people passively browse
- Web is a place where people:
 - Act as prosumers, i.e., content producers and content consumers
 - Interact with other people:
 - Internet forums, Blogs, Social networks, Twitter, Wikis, Podcasts, Slide sharing, Bookmark sharing, Product reviews, Comments, ...

Web and Social Media: Opportunities

- Any user can share and contribute content, express opinions, link to others
- This means: Can data-mine opinions and behaviors of millions of users to gain insights into:
 - Human behavior
 - Marketing analytics
 - Product sentiment

Applications: Real-time citizen journalism

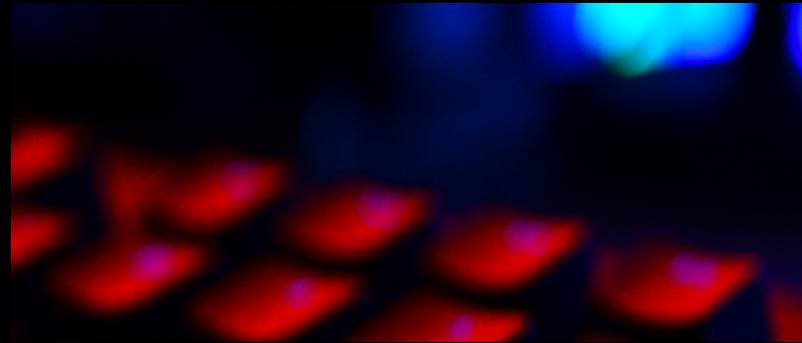
The collage includes:

- A Twitter page for "MumbaiAttacks" (@MumbaiAttacks) with a message: "Hey there! MumbaiAttacks is using Twitter. Twitter is a free service that lets you keep in touch with people through the exchange of quick, frequent answers to one simple question: What are you doing? Join today to start receiving MumbaiAttacks's updates."
- A screenshot of the U.S. State Department's website with a link to "Mumbai Terror Attacks 2008".
- A screenshot of THE TIMES OF INDIA website with a headline: "Real Time Citizen Journalism in Mumbai Terrorist Attacks". It includes a quote: "The Mumbai terror attack has finally ended after more than 60 hours. Even as I continue to track instances of citizen journalism in the Mumbai terror attack on this post, I'm trying to make sense of what happened in a [work-in-progress case study](#) and a [Flickr set of screenshots](#) on the role of social media in the Mumbai terror attack. I'm also compiling [reactions on Indian news media's coverage of the terror attack](#). For more, see my interviews on the role of citizen journalism in the terror attack with [Los Angeles Times](#), [CBS News](#), [BBC DNA](#), [LiveMint](#), [Associated Press](#), [Journalism.co.uk](#), [Teherka](#), [NPR](#), [CNN](#), [CUNY TV](#) and [Star Telegram](#).
- A screenshot of THE TIMES OF INDIA website with a headline: "Terror attacks in Mumbai; six foreigners among 101 dead".
- A screenshot of ALJAZEERA.NET website with a headline: "Scores killed in Mumbai attacks". It includes a photograph of a scene at a railway station.

- **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 the web



On the web...

Anyone can use multiple identities
to put something on the internet.

From anywhere in the world.

They can say anything they like.

Leave it there as long as they like.

Change it whenever they feel like it.

Not everyone has good intentions...

How Trolls Are Ruining the Internet

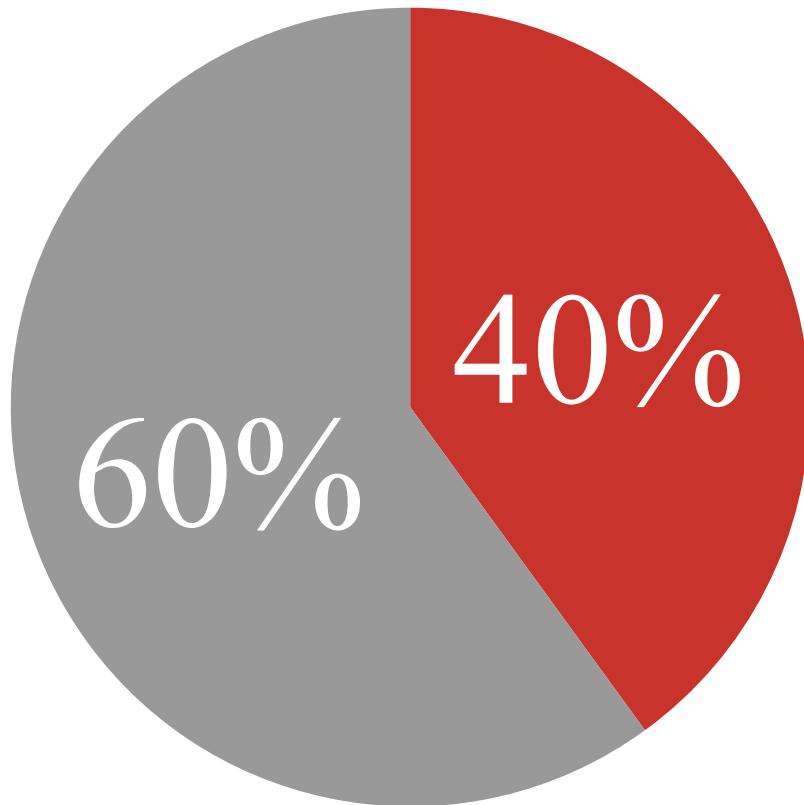
Wikipedia blocks hundreds of 'scam' sock puppet accounts

When Will the Internet Be Safe for Women?

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

Fake reviews on the Play Store reportedly growing and getting smarter

Malicious behavior is prevalent online



40% people
have reported
to have been
harassed online

Malicious behavior is contagious

Trolling behavior increases after seeing already existing troll posts. ($p < 0.05$ using a mixed effects logistic regression model)

% Troll posts	Positive mood	Negative mood
Positive context	35%	49%
Negative context	47%	68%

Anonymity exacerbates the problem



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

©The New Yorker Collection 1993 Peter Steiner
From cartoonbank.com. All rights reserved.

Malicious behavior has harmful effects



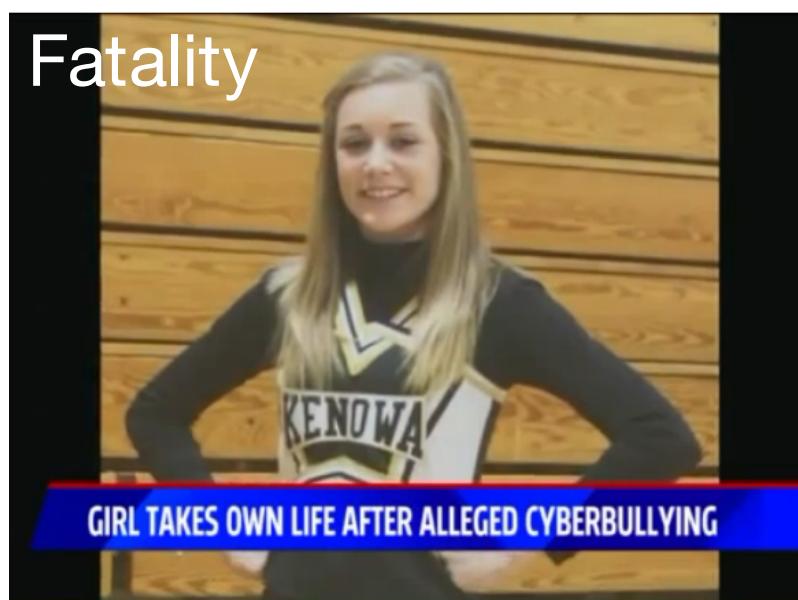
Distress



Harassment



Delinquent behavior



Akbulut et al. (2010), Wiener (1998), Hinduja and Patchin (2007, 2010)

Not everything on the web is 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.

Source: People.com

<http://www.people.com/people/article/0,,20477260,00.html>

威龍 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

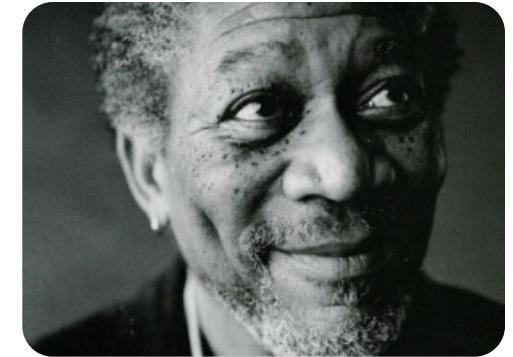
Not everything on the web is true...

- Morgan Freeman died in December 2010 and the news was credited to CNN... according to numerous “tweets” on Twitter.

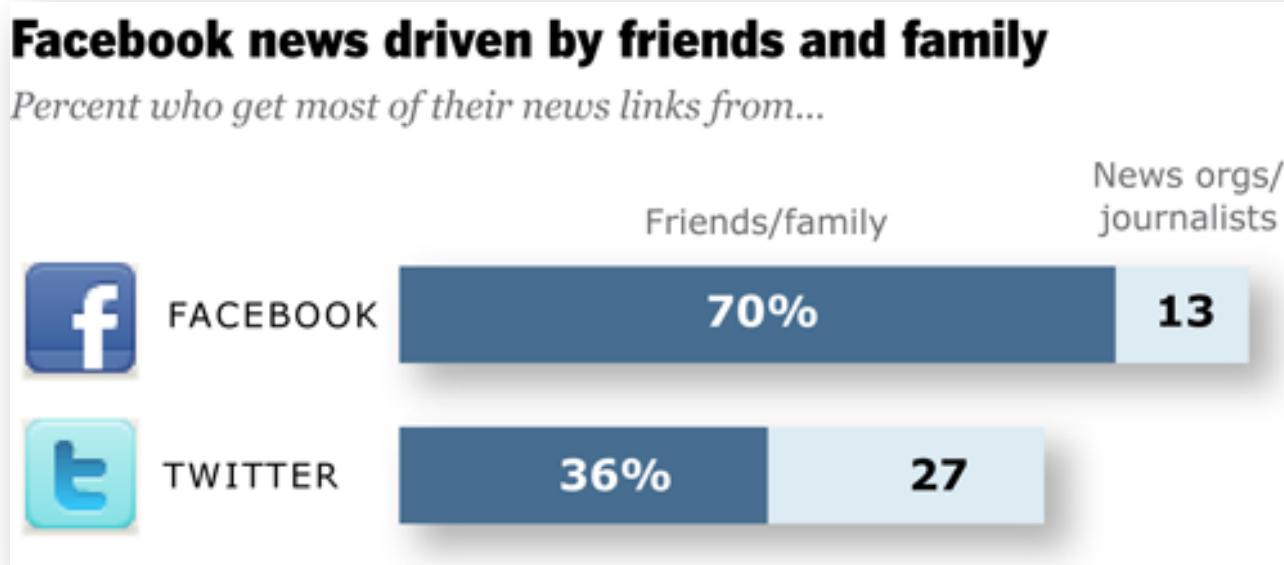
CNN then had to clarify that they never reported the news and that Morgan Freeman was, in fact, alive.

Source: Mediate

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



Misinformation goes viral as people rely on social links for information



Graphic Source: Pew Research Center's Project for Excellence in Journalism / 2012 State of the News Media

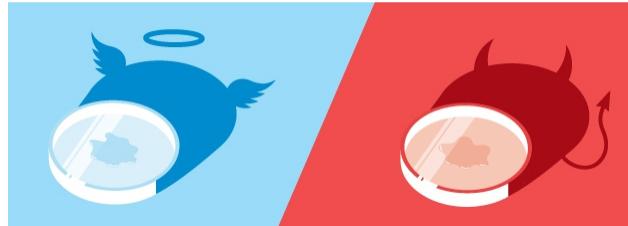
This tutorial:
Modeling and Detection of
Malicious Behavior on the
Web

Challenges in analyzing malicious behavior



Data imbalance

Smaller proportion of behavior (< 10%) is malicious



Limited labels

Little known information about malicious behavior



Deceptive behavior

Malicious behavior tends to masquerade as benign

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

Overarching takeaway

Malicious behavior differs from benign behavior both with respect to their properties and connectivity.

These differences can be used to create efficient computational methods to detect them.

Outline

Introduction

Feature-based algorithms

Sockpuppets

Vandals

Hoaxes

Spectral-based algorithms

Visualization: “spokes”, “blocks”, “staircases”

Camouflage

Theoretical guarantee

Density-based algorithms

Ill-gotten Likes

Synchronized Behaviors

Advertising campaigns

Social spam

Conclusions and future directions

Tutorial link: <http://bit.ly/kdd2017>

Acknowledgement



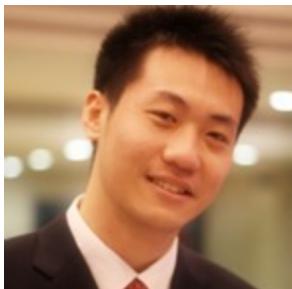
腾讯
Tencent



KDD 2017

Halifax, Nova Scotia - Canada
August 13 - 17, 2017

Tutorial: Data-Driven Approaches towards Malicious Behavior Modeling



Meng Jiang
University of Notre Dame



Srijan Kumar
Stanford University



Christos Faloutsos
Carnegie Mellon University



V.S. Subrahmanian
University of Maryland, College Park

Tutorial link: <http://bit.ly/kdd2017>