The spread of misinformation in social media

Filippo Menczer
Center for Complex Networks and Systems Research
School of Informatics and Computing
Indiana University, Bloomington
Bossy Cracka @Junebug1952 · Sep 25
Obama Forbids FBI to use Religion in Identifying Terror Threats as ISIS OPENLY recruits in Mosques In.is/www.breitbart… Obama IS a MUSLIM

Tara McClintick @BooksByTara · 6h
Thank GOD for "internet information" bc mainstream media hides what thousands know: Vaccines CAUSE autism #CDCwhistleblower

Maggie Hittie @MHittie · Nov 12
WE PAY 4 THIS! @SteveWorks4You: More targeting: Obama spent $1M 4 project 2 suspend conservatives' TwitterAccounts washingtonexaminer.com/rep.-lamar-smi
1. Detection of misinformation
2. How misinformation spreads
3. Can the spread of misinformation be mitigated?
Observatory on Social Media

This website highlights the results of a broad research project aimed to study information diffusion in social media. Use our tools to explore how people spread ideas through online social networks.

Recent Papers from @TruthyAtIndiana

Truthy @truthyatindiana
Hoaxy: A Platform for Tracking Online Misinformation #truthypaper @www2016ca @snow_workshop cnets.indiana.edu/blog/2016/03/0...

Recent Talks from @TruthyAtIndiana

Congrats @clayadavis for best presenter prize at #www2016 Dev Day for truthytalk on @TruthyBotOrNot via @jahendler
Tools
The who, what, where, when, and how of social media data

BotOrNot
Check how bot-like a Twitter user behaves.

Trends
Compare when memes gain and lose popularity.

Movies
Generate movies of how conversations about a meme evolve over time.

Networks
Explore who is discussing a meme and what memes are related.

Maps
Examine where people are talking about a meme over time.

API
Query Truthy's data for your own analysis.
Coming Soon
co-occurrence of #kittens between 2016-03-27T04:00:00 & 2016-03-30T04:00:00

EPJ Data Science 2014
Social Media Observatory

Middleware

NoSQL Distrib. DB

Analytics

Data Collection System

Long-term Backup

Dynamic Vis Algorithm

API

Timeline

Network Vis

Geo Maps

Videos

Dynamic Vis

Algorithm
Here you find our collection of Memes (#hashtag, @user, http://url, or "phrase") from the past 90 days of Twitter communication.

For each meme, we calculate statistics and provide interactive interfaces that visualize the networks, allow you to identify interesting users, and download data.

For each meme we show here a user-generated definition from tagdef.com, a timeline of collected tweets, and the static diffusion network.

<table>
<thead>
<tr>
<th>Meme</th>
<th>Timeline</th>
<th>Diffusion Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>#bahrain</td>
<td>Middle East Country</td>
<td>Tweets</td>
</tr>
<tr>
<td>#syria</td>
<td>A country in the Middle East that has fought throughout the times for its freedom. Lately from its d...</td>
<td>Tweets</td>
</tr>
<tr>
<td>#egypt</td>
<td>An Arabic country in North Africa</td>
<td>Tweets</td>
</tr>
</tbody>
</table>
Static Network

This static communication network shows how the meme is transmitted from user to user, by means of retweets (forwards) and mentions (direct messages).

LEGEND
User
Retweet
@Mention
Astroturf Detection

<table>
<thead>
<tr>
<th>nodes</th>
<th>Number of nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>edges</td>
<td>Number of edges</td>
</tr>
<tr>
<td>mean_k</td>
<td>Mean degree</td>
</tr>
<tr>
<td>mean_s</td>
<td>Mean strength</td>
</tr>
<tr>
<td>mean_w</td>
<td>Mean edge weight in largest connected component</td>
</tr>
<tr>
<td>max_k(i,o)</td>
<td></td>
</tr>
<tr>
<td>max_k(i,o)_user</td>
<td></td>
</tr>
<tr>
<td>max_s(i,o)</td>
<td></td>
</tr>
<tr>
<td>max_s(i,o)_user</td>
<td></td>
</tr>
<tr>
<td>std_k(i,o)</td>
<td></td>
</tr>
<tr>
<td>std_s(i,o)</td>
<td></td>
</tr>
<tr>
<td>skew_k(i,o)</td>
<td></td>
</tr>
<tr>
<td>skew_s(i,o)</td>
<td></td>
</tr>
<tr>
<td>mean_cc</td>
<td></td>
</tr>
<tr>
<td>max_cc</td>
<td></td>
</tr>
<tr>
<td>entry_nodes</td>
<td>Number of unique injections</td>
</tr>
<tr>
<td>num_truthy</td>
<td>Number of times ‘truthy’ button was clicked for the meme</td>
</tr>
<tr>
<td>sentiment scores</td>
<td></td>
</tr>
</tbody>
</table>

### Classifier Performance

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdaBoost</td>
<td>96.4%</td>
<td>0.99</td>
</tr>
<tr>
<td>SVM</td>
<td>95.6%</td>
<td>0.95</td>
</tr>
</tbody>
</table>

The size of the largest connected component
Number of unique injections
The six GPOEMS sentiment dimensions
ICWSM 2011

The Curious Case of Cynk, an Abandoned Tech Company Now Worth $5 Billion

Pols have a #fakefollower problem

The Fembots of Ashley Madison
Real-time analysis and classification

Real-time query (Twitter search API)
Real-time feature (>1K) extraction
Real-time analysis and classification

truthy.indiana.edu/botornot
AUC 0.95
BotOrNot Python API

A Python API for Truthy BotOrNot.

Behind the scenes, this uses the BotOrNot's HTTP endpoint as illustrated in this notebook.

Quickstart

```python
import botornot

twitter_app_auth = {
    'consumer_key': 'xxxxxxxxx',
    'consumer_secret': 'xxxxxxxxxx',
    'access_token': 'xxxxxxxxxx',
    'access_token_secret': 'xxxxxxxxxxxxx',
}

bon = botornot.BotOrNot(**twitter_app_auth)

bon.check_account('@clayadavis')
```
DARPA Twitter bot detection challenge

<table>
<thead>
<tr>
<th>Team</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentimetrix</td>
<td>50.75</td>
</tr>
<tr>
<td>USC</td>
<td>45.00</td>
</tr>
<tr>
<td>IU</td>
<td>43.25</td>
</tr>
<tr>
<td>IBM</td>
<td>43.00</td>
</tr>
<tr>
<td>Boston Fusion</td>
<td>41.75</td>
</tr>
<tr>
<td>Georgia Tech</td>
<td>24.00</td>
</tr>
</tbody>
</table>

*IEEE Computer, June 2016*
1. Detection of misinformation

2. How misinformation spreads (Truthy case study)

3. Can the spread of misinformation be mitigated?
Concerned American @ObamasMentors · Aug 26
Feds Spend $1 Million to track "hate speech" on twitter, run by a FAR LEFT LIBERAL-An OBVIOUS ATTACK ON CONSERVATIVES-freebeacon.com/issues/feds-cr...

Youxia88™ @Youxia88 · Aug 27
@jjauthor Only conservatives will be targeted. RT @FoxNews: Feds creating database to track hate speech on Twitter fxn.ws/1qoNlz9

f. lindsay @hidefullofHoney · Aug 28
#Truthy: How Feds Plan to Track Hate Speech (Conservatives) on Twitter (TWTR) @CNN @msnbc @ABCNews @CBSNews @NBCNews invstplc.com/1nlDkO4
Dean R. Riehm @deanriehm · Oct 21
dailycaller.com/2014/10/21/the... Oh, look who’s back: it’s those taxpayer-funded fascist anti-free speech thugs. #truthy @truthy @BDaddyLiberator

Leviathan Rising @govtblackhole · Oct 23
Big #Gov targets conservative #speech on Twitter. wapo.st/ZyAi9c #freespeech #tlot #tcot #pjnet #teaparty #censorship #ccot #law

RockPrincess @Rockprincess818 · Oct 27
Secret Police. Obama’s regime is funneling our tax dollars to radical anti-free speech groups using public services as his bag man #Truthy

#Truthy, Because Universities Should Encourage The Free Exchange Of Ideas
THE Arizona Luke @ArizonaLuke · Nov 10
Govt Funded "Truthy Project" Boasts That It Suspended #tcot Twitter Accounts. #Censorship #FreeSpeech freebeacon.com/issues/house-c...

Laura LaRue @LaLaRueFrench75 · Nov 10
@NSF Why are you having #Twitter accts suspended for #FreeSpeech & using my tax $ to do it? freebeacon.com/issues/house-c... ow.ly/i/7xdlH

Mary Forbes @MaryForbes14 · Nov 11
ICYMI Obama's Administration paid twitter trolls to target Conservatives for Twitter suspension drudge.tw/1uUSYy2

Chuck DeVore @chuckdevore · Nov 11
Oh, this is great, using almost $1 million in tax $ to suspend conservative Twitter accounts. freebeacon.com/issues/house-c... #tcot

Red Alert Politics @RedAlert · Nov 11
"Taxpayers deserve to know" how NSF decided to award a large grant to project suppressing free speech redalert.io/1tYqqRo

GENUG @GENUG · Nov 15
Big Truthy Is Watching (Some of) You - ammoland.com/2014/11/big-tr... via @Ammoland #Obama’s #fascist #censorship #project #monitoring #free #speech

Daniel Shoop @Shoopdj · Jan 17
Indiana University Monitored ‘Hate Tweets’ in Aftermath of Truthy Scandal ln.is/freebeacon.com... The Fed’s version of “Free Speech”.

Laura Van Overscheld @LauraVanOversch · Dec 27
House Committee Demands Answers on Truthy Project ln.is/freebeacon.com... #Filippo Menczer-self-appointed-free-speech-monitor-on-Twitter

Silverbelle1008 @silverbelle1008 · Nov 11
Scary lib out to shut down conservative free speech House Committee Demands Answers on Truthy Project freebeacon.com/issues/house-c...
La ciencia y la tecnología puestas al servicio de la persecución contra la libertad de expresión: EE.UU. lanza una guerra millonaria contra los “memes sospechosos”

En EE.UU. se ha desarrollado un software que permite contrarrestar la “propaganda” y las “incitaciones al odio” en las redes sociales. Sin embargo, todos sabemos realmente los lamentables fines del desarrollo de este tipo de tecnologías.
Competition between hoaxes and fact checking

chemtrails

anti-vax
Hoax vs. fact checking: model

\[ f_i = f = \frac{n^{AB}(i)}{n(i)} \]

- **Susceptible**
  - \( n(i) \) #neighbors of node i
  - \( \alpha \) credibility
  - \( p_v \) verifying probability
  - \( \tau^{-1} \) decay

- **Believer**
  - \( \tau^{-1} \)
  - \( f \)

- **Fact-Checker**
  - \( \tau^{-1} \)
  - \( f \)

- **Model Diagram**

- \( \alpha f \)
- \( (1-\alpha) f \)
- \( p_v \)
Hoaxy

Social Networks
- Facebook
- Twitter
- Instagram
- LinkedIn

News Sites
- HoaxAndLegit
- Snopes

API
- URL Tracker (stream api)

Crawler
- RSS Parser
- Scrapy Spider

Monitors

Store

DATABASE

Fetch

Analysis Dashboard
1. Detection of misinformation
2. How misinformation spreads
3. Can the spread of misinformation be mitigated?
COMPUTATIONAL FACT-CHECKING?
is the spouse of

Woodrow Wilson
Warren G. Harding
Calvin Coolidge
Herbert Hoover
Franklin D. Roosevelt
Harry S. Truman
Dwight D. Eisenhower
John F. Kennedy
Lyndon B. Johnson
Richard Nixon
Gerald Ford
Jimmy Carter
Ronald Reagan
George H. W. Bush
Bill Clinton
George W. Bush
Barack Obama

is the capital of
Bossy Cracka @Junebug1952 · Sep 25
Obama Forbids FBI to use Religion in Identifying Terror Threats as ISIS OPENLY recruits in Mosques in.is/www.breitbart... Obama IS a MUSLIM

Kathryn Master @KathrynMaster1 · Sep 24
@megynkelly #Obama is a Muslim who wants to turn America into Muslim Brotherhood controlled country. Pls research & do show on. Must stop him

mary christie @peaceischrist · Sep 24
Husain Barack Obama is a Muslim! That is why he hates America, Christians, Israel etc! #Obama @BarackObama @oreillyfactor
Knowledge Linker

Subject: Bill Clinton  
Predicate: spouse  
Object: Hillary Rodham Clinton

Most Informative Path:

Bill Clinton ➔ Chelsea Clinton ➔ Hillary Clinton
DOES FACT-CHECKING WORK?

- Echo chambers
- Selective exposure
- Confirmation bias
Predicting political alignment

<table>
<thead>
<tr>
<th>Features</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text (TF-IDF)</td>
<td>79%</td>
</tr>
<tr>
<td>Hashtags</td>
<td>91%</td>
</tr>
<tr>
<td>Retweet network</td>
<td>95%</td>
</tr>
<tr>
<td>Tags + Network</td>
<td>95%</td>
</tr>
</tbody>
</table>

SocialCom 2011
Social Scientists Wade Into the Tweet Stream

*Science* 30 September 2011:
vol. 333 no. 6051 1814–1815

**Echo chamber.**

On Twitter, liberals (blue) and conservatives (red) network more with like-minded users, with less traffic crossing the political divide.

CREDIT: © 2011 TRUTHY PROJECT (TRUTHY.INDIANA.EDU), INDIANA UNIVERSITY

**Activity**

**K-core follower network**

**K-core retweet network**
Infographic Of The Day: Does Twitter Make Our Politics Worse?

By Cliff Kuang Thu Jul 28, 2011
Homogeneous Exposure

Social bubbles

Entropy

Social Media | Search

PeerJ CS 2015
COMPETITION FOR ATTENTION

Hashtag Popularity

# daily retweets
[Twitter]

User Popularity

# followers
[Yahoo! Meme]
Can the competition for limited user attention help explain the broad heterogeneity of meme popularity and our vulnerability to misinformation?
Toy agent-based model

Before

Post a new topic (Pn)

Post existing topics (1 - Pn)

Screen (Pr)

#apple
#jobs
#justinbieber
#ladygaga

Memory (Pm)

#apple
#jan25

After

Screen (Pr)

#apple
#jan25
#jobs

Memory (Pm)

#jan25
#apple
#jobs
#justinbieber
#ladygaga
Toy model predictions: Role of social network

![Graphs showing meme lifetime and popularity distributions.](image)
Toy model predictions: Role of \textit{limited attention}
• Spread among agents with **limited attention on social network** is **sufficient** to explain meme virality

• *Not necessary* to invoke more complicated explanations based on intrinsic meme value or external factors
DO THE BEST IDEAS WIN?
LIMITED ATTENTION

![Graph showing diversity vs. efficiency with markers for different α values.](image)
• Structural, temporal, content, and user features can be used to detect astroturf and social bots.

• Social media and traditional media work together to spread misinformation; gullible people are exposed through gullible connections.

• Social network structure and limited attention may amplify our natural biases and make us more vulnerable to misinformation.
Thanks!