Introduction to Fake News

Dongwon Lee

dongwon@psu.edu

Penn State University, USA
Fake News is Not New

**GOVERNMENT COVER-UP EXPOSED!**

THE DOWNING OF TWA FLIGHT 800

THE SHOCKING TRUTH BEHIND THE WORST AIRPLANE DISASTER IN U.S. HISTORY

JAMES SANDERS

---

**Hundreds of Eyewitnesses Know the FBI and CIA Lied!**

We Saw TWA Flight 800 Shot Down by Missiles

And We Won’t Be Silenced Any Longer

---

**Feds deny claim that Navy shot down TWA 800**

But investigators have said on several occasions there is no evidence to support the friendly fire theory.

And again in a joint statement Friday, the FBI, Navy and the National Transportation Safety Board said:

“The so-called ‘friendly fire’ explanation has yet to be proven by the U.S. government.”

There is no evidence to support the friendly fire theory.

The 71-year-old Sawyer showed reporters his book on two crumpled pages, but wouldn’t let them read it. He said it was dated Aug. 31, and given to him five weeks ago by someone in French intelligence, but written by an American who “was tied to the U.S. Secret Service and has important contacts in the U.S. Navy.”

Lt. Cmdr. Rob Newell, a Navy spokesman at the Pentagon, said the nearest warship, the USS Normandy, an Aegis-type missile cruiser, was 166 miles to the south and was not conducting weapons drills. Its computer-driven, long-range radar system was on “low power,” engaging only 150 miles, and “could not even see the TWA plane.”

Newell said the only aircraft in the area was a P-3 Orion anti-submarine plane, which he said does not carry missiles. But the military reference, James All The World’s Aircraft, states the P-3 is capable of carrying missiles.

---

very important people,” to wait until after the U.S. election to disclose the information because it could “ruin the public’s very carefully built-up

The FBI has enjoyed doses of calls from people claiming to have seen a missile at the time of the explosion. Rumors that the aircraft was accidentally shot down by the military have abounded on the Internet.
Definitions of “Fake News”

- Fake news was exclusively used in the satire context
  - Eg, The Daily Show with Jon Stewart
  - Fake news is meant to perceived as unrealistic, while traditional news content is meant to be perceived as realistic [Balmas, 2014]

- News articles that are intentionally and verifiably false, and could mislead readers [Allcott et al, 2017]

- Fake news vs. alt-facts: Alt- facts are different from fake news in that they have no basis in reality
  - Statements are evaluated based on “goodness of fit with a particular ideology, but not facts” [Berghel, 2017]

- Journalism vs. fake news: Gatekeeping—“attempt at exercising reliability, selecting the important over the trivial while avoiding sensationalism” [Bordem et al. 2017]
 Definitions of “Fake News”

- “The task of categorizing news along a continuum of veracity [intentional deception], with an associated measure of certainty” [Conroy et al., 2015]
- Intentionally false and shared via the Internet [Klein et al., 2017]
- “Everything from malicious stories to political propaganda” [Cohen, 2017]
  - Many articles are written by journalists who write articles using web searches but, with no actual verification [Cohen, 2017]
  - 53.8% of journalists use microblogs (ex. Twitter) to gather information and report from news stories [Willnat, 2014]
- Fake news as a continuum because there is some news that have a biases and do not fall directly into fake news [Novotny, 2017]
Related Terms

- Misinformation: false information
- Disinformation: deliberately false information
- Rumor: unverified information at the time

https://www.kdnuggets.com/2016/08/misinformation-key-terms-explained.html
So, What is “Fake News”?

I think it's a disgrace that information that was false and fake and never happened got released to the public.

https://www.youtube.com/watch?v=wJxxQM7GxJA
Surge of “Fake News”: Google Trend

Related queries

1. trump fake news
2. trump
3. trump news
4. facebook fake news
5. fake news awards
# Surge of “Fake News”: Google Trend

<table>
<thead>
<tr>
<th>Rank</th>
<th>Country</th>
<th>Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>United States</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>Philippines</td>
<td>92</td>
</tr>
<tr>
<td>3</td>
<td>Singapore</td>
<td>77</td>
</tr>
<tr>
<td>4</td>
<td>Canada</td>
<td>66</td>
</tr>
<tr>
<td>5</td>
<td>United Kingdom</td>
<td>66</td>
</tr>
</tbody>
</table>
Impact of “Fake News” in 2016

Total Facebook Engagements for Top 20 Election Stories

ENGAGEMENT REFERS TO THE TOTAL NUMBER OF SHARES, REACTIONS, AND COMMENTS FOR A PIECE OF CONTENT ON FACEBOOK SOURCE: FACEBOOK DATA VIA BUZZSUMO

Change: Role of ”Receivers”

- Interactive Media ➔ fundamental shift in communication
- Receiver becomes the new “sources”

[Shannon and Weaver, 1949; Berlo, 1960; Sundar and Nass, 2001]
Change: Self as Creators and Distributors
Inside a Fake News Sausage Factory: ‘This Is All About Income’

By ANDREW HIGGINS, MIKE McIntire and GABRIEL J.X. DANCE   NOV. 25, 2016

A fake news article posted by Beqa Latsabidze, a Georgian. “My audience likes Trump,” he said. “I don’t want to write bad things about Trump. If I write fake stories about Trump, I lose my audience.”

https://goo.gl/seKndN @ NY Times
The reality of source layering:

– Imagine checking your Facebook news feed and seeing something your friend has shared, a politician’s tweet of a newspaper story

– Who or What is the Source? Newspaper, Politician, Twitter, Friend or Facebook?
Studies show that only the proximate source matters

Increasingly, these proximate sources are not news organizations, but

– Aggregator sites (eg, Comcast, Verizon)
– Social networking sites (where your friends are sources)
Credit: Hoaxy @ Indiana U.
Complex Issue to Solve

- Conspiracy Theories
- Confirmation Bias
- Selective Exposure
- Selective Perception
- Filter Bubbles
- Internet Credibility
- Media Literacy
- Information Literacy
- Research
- Education
Types of Fake News

Did an Alabama Church Display a Sign Comparing Roy Moore to Jesus?
5 December 2017 - Living Way Ministries said its pastor did not approve of the sign, but at least two members of the congregation posted photos of it on social media.

Did Tokyo Open the First Human Meat Restaurant?
2 December 2017 - Reports that a new restaurant recently opened in Japan is the first in the world to serve meat from human corpses.

Does the 2017 Tax Reform Bill Eliminate Small Business Deductions?
4 December 2017 - Contrary to rumor, the tax bill does not eliminate deductions for small business expenses, but rather changes the rules for deducting employee compensation.

Does the 2017 Republican Tax Bill Cut Cancer Treatments for Medicare Patients?
4 December 2017 - The Tax Cuts and Jobs Act does not mention the healthcare program, but critics say cancer patients would be hit hard if the bill passes.

Massive Indian Ocean Earthquake and Tsunami Expected by the End of 2017?
27 September 2017 - An Indian man’s prediction is based on his purported psychic abilities and carries no scientific weight.

Did a CIA Agent Confess to Killing Bob Marley?
4 December 2017 - A disreputable website capitalized on an old conspiracy theory about the death of the reggae icon to spin fake news clickbait.
• “Attempt at exercising reliability, selecting the important over the trivial while avoiding sensationalism” (Bordem & Tew, 2007, p. 304).

• “The Opinion Journalist has permission and is licensed to express opinions. His/her job is to select facts to form an argument” (Digital Resource Center, 2017)

Lexical and Syntactic structure: Use of past tense reporting

Reports from “official sources”

Adheres to AP style

Labeled as “opinion”

Lexical use: “should”

Emotionally charged linguistic markers

Real News

Commentary
Satire

A story [intentionally false] generated in a non-professional social context that uses the style of news either to parody that style, satirize issues and personalities in the news, or perpetrate a hoax or prank. (Frank, 2015, p. 317).

Citizen Journalism

Blogs, websites, and posts dedicated to citizen journalism, citizen generated petitions, personal activism, and other forms of civic expression that display originality and creation more than curation or aggregation (Howard et al., 2017).

Trump Comforts Grieving War Widow By Assuring Her He Will Never Die

WASHINGTON—In an effort to alleviate her sorrow after the tragic event, President Donald Trump reportedly comforted the grieving widow of a fallen soldier Thursday by assuring her that he will never die. “I know things are very difficult for you right now, but I hope you can take some solace in the knowledge that I am going to live on forever,” said Trump in a phone call to Deborah Stilling, attempting to soothe the woman who lost her spouse in Afghanistan with promises that he would outlast her, her children, and her children’s children. “Yes, your husband is gone, but remember that the commander-in-chief served will always remain, and no one can ever take that away from you. Dry your eyes; I’m still here.” At press time, Trump, who was reportedly unsuccessful in lifting the widow’s spirits, offered repeated assurances that he would thrive long after the sun had expanded into a red giant and swallowed the Earth.

W300.png

Clickbait headline

About us section

Amateur mistakes

Not fact-checked

URL and about us

Might not follow journalistic practices

Can be emotionally charged
Eg, Native Ad. & Professional Political Content

Native Ad.

Promotional material disguised as news article (Frank, 2015)

Information from government, political parties, or public agencies.

No sources, or sources related to the brand

Paid by brand: logo or name of company

Lexical used: present tense

Prof. Political Content

Lack of source, or sources are from members of the group

URL

About us

No sources, or sources related to the brand

Paid by brand: logo or name of company

Lexical used: present tense
Polarizing/Sensational Contents

- Attention grabbing techniques
- Emotionally charged
- Absence of quotes
- Lexical structure: first person

Fake Sites

- Emotional linguistic markers, and attention grabbing techniques
- Ephemeral
- Structural inconsistencies: URL registration

Example:

**Polarizing/Sensational Contents**

- "What happened?" That is the million dollar question. Clinton has bought to the UK this week. And after the interviews, I’m still none the wiser as to what...

Hurricane Harvey

**Fake Sites**

- "Las Vegas Shooter Reportedly a Democrat Who Liked Rachel Maddow, MoveOn.org and Associated with Anti-Trump Army"

Ephemerality: URL registration
Fake news? Not if 2 PSU profs can help it
Penn State Professors Want To Build Device That Recognizes Fake News

By ASSOCIATED PRESS  •  AUG 31, 2017

A pair of professors from Penn State University received a $300,000 grant in an effort to help produce technology — specifically “devices” — that would be able to efficiently detect fake news, according to the Associated Press.
The university says information sciences and technology professor Dongwon Lee and communications professor S. Shyam Sundar are working on the project.

Lee says fake news "has been around for decades" but has been "exacerbated" on the Internet and social media platforms.

The professors plan to investigate "characteristic indicators of fake news" and develop complex formulas that will enable digital devices to recognize those indicators and purge stories that contain them.
Lol, a bullshit detector. This is the one thing that keeps our ears from getting blown off, that's right, it's called critical thinking and being able to separate fact from fiction, programming, and regulation. Seems to be an art lost to communities like Breitbart and InfoWars.

They could start by looking into the story of how a fraternity brother died while on a mission trip to the Dominican Republic.

Another step to regulate free speech and shape the narrative/brainwash the public.
Important Questions

- **What is Fake News?**
  - Multi-faceted, Multi-types
  - Address more impactful fake news facet/type
    - Eg, Biased/opinionated stories are not urgent

- **How to know if solution is good enough?**
  - Accuracy threshold: Eg, smarter than human?
  - Trust issue

- **Virality as key factor**
  - Traditional media research looks for *intention* and credibility
    - No longer simple in internet era
  - How to make people think twice before re-posting fake news?
  - How to label it?
Important Computational Questions

- How does fake news spread?
  – Understand the propagation mechanism of fake news

- Who are the users susceptible to fake news?
  – Who are creators or conveyors of fake news?

- Detect fake news among real news
  – Detect it early
Related Detection Problems in Literature

- Fake News Detection
- Clickbait Detection
- Rumor Detection
- Truth Discovery
- Social Bot Detection
- Social Spammer Detection
Fake News Detection: Human-Oriented

- White/blacklist-Based
  - BS detector: http://bsdetector.tech/
  - Rubtr: http://rbutr.com/
  - ZenMate SafeSearch: https://zenmate.com/products/safesearch/
  - FiB: https://devpost.com/software/fib
  - Media Bias/Fact Check: https://mediabiasfactcheck.com/

- Manual fact-checking

- FactCheck.org
- PolitiFact.com
- OpenSecrets.org
- Snopes.com
- TruthOrFiction.com
- Hoax-Slayer
Computational: AI vs. DB Approaches

- **AI Approach**
  - Underlying premise: the existence of latent patterns to differentiate fake news from real news
  - Costly to develop labeled samples
  - Can handle fresh fake news better

- **DB Approach**
  - Underlying premise: if query (= news) to DB/KB of facts yields null, it is fake news
    - Closed-world assumption (CWA)
  - Costly to build/maintain DB/KB
  - May not handle fresh fake news
  - Can explain better
Online Quiz: Fake-O-Meter

- In your browser, go to **Kahoot.it**
- Enter Game PIN, shown on screen
- Enter your nickname (to be shown on screen)
Landscape of Machine Learning Approaches

- **Unsupervised Learning**
  - Fake news \(\subseteq\) outlier
  - Plenty of *unlabeled* dataset
  - Lower detection accuracy

- **Supervised Learning**
  - Scarce and unbalanced *labeled* dataset
  - Hand-crafted feature engineering

- **Deep Learning**
  - Difficult to explain
  - Post-detection issue
Supervised Approaches

Training

- Supervised Machine Learning
  - \( P \): Features from "fake" news
  - \( N \): Features from "legit" news
- Feed \((P, N)\) to ML to build a model \( M \)

Testing

- Feed a news story \( A \) to \( M \)
- \( M \) determines if \( A \) is fake or legit news story
Features: Fake News Example

Latest Dead Body Tied To The Clintons Turns Up Valuable Evidence

December 20, 2017 Flagg Eagleton Loose Meat Hoagies, Potater Special

Another body found in a shallow grave in Arkansas has been tied to the Clintons. Marcia Ann Bosley, who was the head housekeeper at the Governor's Mansion before she was mysteriously let go in 1985, has been positively identified after spending more than 3 decades in the ground.

Bosley, who had no family when she was let go, was assumed to have left the area for other opportunities. She was never reported missing. According to the Hamilton County Coroner:

About Flagg Eagleton

Flagg Eagleton is the son of an American potato farmer and a patriot. After spending 4 years in the Navy and 7 on welfare picking himself up by the bootstraps, Flagg finally got his HVAC certificate and is hard at work keeping the mobile homes of Tallahassee at a comfy 83 degrees.

https://www.facebook.com/thelastlineofdefense/
Features: Fake News Example

Donald J. Trump
@realDonaldTrump

#CrookedHillary thinks she can win in Utah with loser Mitt Romney’s help. Who cares about Utah? Small state, strange people. Boring place.

983 RETWEETS 2155 FAVORITES

11:34 AM - 10 Aug 2016
<table>
<thead>
<tr>
<th>Features: Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fake News</strong></td>
</tr>
<tr>
<td><strong>Lexical</strong></td>
</tr>
<tr>
<td><strong>Syntactic</strong></td>
</tr>
</tbody>
</table>
| **Message quality** | • Spelling mistakes  
• Not in AP style  
• No last names when citing | • Goes through proof  
• Reading and editing process  
• Adheres to AP style |
| **Pedigree** | Originated in an obscure site or social media post | Originated from a well-known site |
| **Quotes** | • Unverified sources  
• Absence of quotes | • Verified quotes (real people and quotes)  
• Always has quotes, if not it is a commentary |
| **Photos** | • Altered pixel structure  
• Picture is out of context with the event | • Normal pixel structure  
• Pictures will correspond to the event at hand |
Credibility

Originality = original tweets vs. retweeted

Role = 
# followers vs. # following

Influence

Barack Obama
@BarackObama
Dad, husband, President, citizen.

Tweets | Following | Followers | Likes | Lists
--- | --- | --- | --- | ---
15.5K | 627K | 96.3M | 10 | 3

Tweets | Tweets & replies | Media
--- | --- | ---
Barack Obama @BarackObama · Oct 24
I’ll let you and @ladygaga handle the singing, and
There’s still time to give: oneamericappeal.org.
BREAKING: Roger Goodell has signed a contract extension to remain the commissioner of the NFL for an additional 5 years.

That's a full 4 years longer than President Trump will be in office, and it means that the injustices that Colin Kaepernick represents can further be addressed.
Fake News Datasets

- PHEME rumor: https://goo.gl/5JXAXt
- NewsVerification: https://goo.gl/jDZJpS
- BuzzFeedNews: https://goo.gl/TUnw2K
- LIAR: https://goo.gl/wkNYhv
- BS Detector: https://github.com/bs-detector/bs-detector
- FakeNewsSites: https://www.kaggle.com/mrisdal/fake-news/data
Educational Fake News Games

- http://factitious.augamestudio.com/
- https://www.fakeittomakeitgame.com/
- https://playfakenews.com/
- https://hoaxy.iuni.iu.edu/fake-news-game/

- http://fakenews.game/
- https://boardgamegeek.com/boardgame/235085/fake-news-or-not
Reference

- [Ciampaglia et al., 2018] Research Challenges of Digital Misinformation: Toward a Trustworthy Web, AI Magazine
- [Jin et al., 2016] News Verification by Exploiting Conflicting Social Viewpoints in Microblogs, AAAI
- [Kwon et al., 2013] Prominent Features of Rumor Propagation in Online Social Media, ICDM
- [Ma et al., 2017] Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning, ACL
- [Ma et al., 2016] Detecting Rumors from Microblogs with Recurrent Neural Networks, IJCAI

[Qazvinian et al., 2011] Rumor has it: Identifying Misinformation in Microblogs, EMNLP

[Shu et al., 2017] Fake News Detection on Social Media: A Data Mining Perspective, KDD Exploration

[Wu and Liu, 2018] Tracing Fake-News Footprints: Characterizing Social Media Messages by How They Propagate, WSDM

[Wu et al., 2017] Mining Misinformation in Social Media: Understanding Its Rampant Spread, Harm, and Intervention, ICDM

[Zhao et al., 2015] Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts, WWW

[Zubiaga et al., 2017] Exploiting Context for Rumour Detection in Social Media, SocInfo

[Zubiaga et al., 2018] Detection and Resolution of Rumours in Social Media: A Survey, CSUR