

Actor, Culture or Structure? A Predictive Model of News Article Popularity

Weiwei Zheng

Introduction

What influences News Popularity?

- Agenda Setting vs. Agenda Building

Media influences what we think.

Transfer of the media agenda is a reciprocal process.

-- Rogers, Dearing, 1988

- **Audience** takes an active role in virtual sphere during the process of information diffusion. **Human autonomy matters.**
- News popularity can be predicted by **crowd manipulation.**

-- *D. Horne and Adali, 2017*

Introduction

What shapes human behavior?

- "Duality of structure" : social practice, which is the principal unit of investigation, has both a **structural** and an **agency**-component.
 - Anthony Giddens
- **Culture** is an important mechanism.
- “A critical site of social action and intervention, where power relations are both established and potentially unsettled.”
 - Stuart Hall
- Headline negativity, subjectivity and overall sentiment influence news popularity.
 - Reis, Julio, etc., 2015

Research Questions

Goal: To Build a Predictive model on News Popularity

- Can active participants/opinion leaders, news content, and network structure altogether predict the popularity of online news article?
- Hypothesis: Yes.

Significance of the Topic:

- 1) Touch on the relationship between network structure and info diffusion (ameliorate communication strategy online).
- 2) Look into the mechanism and process of developing online public opinion.

Informal Model

Dependent Variable – News popularity

an aggregate of communication effect and public opinion

Predictive Variables – three dimensions

- Structure
Audience Network Attributes
- Agency
Opinion Leaders (number of actors of different roles)
- Culture
News articles' content, sentiment, media agency

Confounders – Date of published, total time of view

Data Sources

- Reddit *an American social news aggregation, web content rating, and discussion website.*
- API: Praw
- /r/news, /r/worldnews (top section)

MY SUBREDDITS ▾ HOME · POPULAR · ALL · RANDOM · USERS · ANNOUNCEMENTS · SOCIOLOGY_ACADEMIC · ASKRREDIT · WORLDNEWS · VIDEOS · FUNNY · TODAYILearned · PICS · GAMING · MOVIES · NEWS · GIFS · MILDLYINTERESTING

Reddit /r/news hot new rising controversial top gilded wiki

Post all analysis/opinion/politics articles to /r/InTheNews

/r/inthenews /r/worldnews /r/politics new comments

links from: past year ▾

Legends aren't born. They're forged on the open seas. Set sail across a shared world to create your own pirate legend and discover what Polygon called "the funniest game we've ever played." (xbox.com)

promoted by Sea_of_Theives

↳ promoted save report

Scientist Stephen Hawking has died aged 76 (news.sky.com)
submitted 20 days ago by araylisusername
185k upvotes 7152 comments share save hide report crosspost

Soft paywall: F.C.C. Announces Plan to Repeal Net Neutrality (nytimes.com)
submitted 4 months ago by IanSimmsMady
178k upvotes 11126 comments share save hide report crosspost

Apple admits it slows older iPhones, confirming Geekbench report (cnet.com)
submitted 3 months ago by ErnestGetsElected
149k upvotes 13499 comments share save hide report crosspost

Soft paywall: Net Neutrality Overturned (nytimes.com)
submitted 3 months ago by DWinsauer
147k upvotes 18855 comments share save hide report crosspost

Japanese firm gives non-smokers extra six days holiday to compensate for cigarette break (independent.co.uk)
submitted 5 months ago by ajstrange1
127k upvotes 5025 comments share save hide report crosspost

Chester Bennington of Linkin Park commits suicide (wtkr.com)
submitted 8 months ago by Another_Chance
124k upvotes 16086 comments share save hide report crosspost

#DeleteFacebook Movement Gains Steam After 50 Million Users Have Data Leaked (thewrap.com)
submitted 14 days ago by JoseTwitterfan
123k upvotes 7134 comments share save hide report crosspost



Data Sources

- News Articles
 - All urls in /r/news (worldnews) top section head title, content, word count
- API: Newspaper3k (supports NLTK)

"All the News That's Fit to Print"

The New York Times

LATE EDITION
Today, Good, afternoon, rainy, high 56. Tonight, a bit of drizzle, low 48. Saturday, mostly cloudy, high 58. Sunday, a late-day shower, high 59. Weather map is on Page 18.

VOL CLXII , No. 56,092 © 2013 The New York Times NEW YORK, SUNDAY, MARCH 31, 2013 \$5.00

As OSHA Emphasizes Safety, Long-Term Health Risks Fester

Toxic Factory Fumes Test Agency's Powers

By IAN URIBA

TAYLORSVILLE, N.C. — Sheri Farley works with a power tool in her only job she could hold would be one where she does not have to stand on her feet for long stretches, otherwise pain screams down her spine and legs. "It's like torture," Ms. Farley describes herself, recalling how she recently had to leave a job she was working at because a doctor was warning to her mother about whether the "crippled lady" was a danger to herself and others.

For about five years, Ms. Farley, 45, stood alongside about a dozen other women, all wearing in-hand, giving together foam cushions for chairs and couches sold by a company called Broyhill, Ralph Lauren and Thomasville. The three companies also formed a yellowish fog inside the plant, and Ms. Farley's doctor told her that the exposure, if it eventually ate away at her nerve endings, resulting in what she calls "nerve" or co-workers called "dead foot."

A chemical she handled —

known as tri-propyl benzene, or TMB — is used by tens of thousands of workers in auto body shops, dry cleaners and furniture manufacturers and tanning plants across the nation. Medical researchers, government officials and environmental companies that once manufactured the chemical warn that over a decade it can cause reproductive damage and infertility when inhaled. It was first used in the early 1940s, but its use has grown 10-fold in the past six years.

For the last two years, Farley has tried to decrease the difficulty, despite the decrease of effort, of ensuring that Americans make safe products for their homes. Even as worker after worker fell ill, records from the Occupational Safety and Health Administration show that managers at Broyhill, Thomasville and Broyhill's sister company, Broyhill Furniture, which Farley was employed, repeatedly exposed glaziers to TMB levels that exceeded federal limits. Managers considered salts, failed to provide respirators and turned off fans to vent fumes.

Photo: AP

SHORT OF MONEY, EGYPT SEEKS CRISIS ON FUEL AND FOOD

ECONOMIC FEARS DEEPEN

Government Is Pressed to Increase Taxes and Cut Subsidies

By DAVID D. KIRKPATRICK

QALYUBIYA, Egypt — A fuel shortage has sent food prices soaring. Electricity is blacking out even before the sun rises. And protesters have killed at least five people and wounded dozens over the past two days.

The root of the crisis, economists say, is that Egypt is running out of money and it needs for fuel imports. The shortage is raising questions about whether the country can keep operating wheat that is essential to subsistence farmers. It is also raising fears of an economic catastrophe at a time when the government is already struggling to quell the protests.

Photo: AP

BYLINE PHOTO BY NICK LAMBERT/CITY PRESS

The Boston Globe

SUNDAY, APRIL 9, 2017

DEPORTATIONS TO BEGIN

President Trump calls for tripling of ICE force; riots continue

Curfews extended in multiple cities

PRESIDENT TRUMP has set in motion one of his most controversial campaign promises, ordering the Justice Department to assemble a "massive deportation force" by tripling the number of immigration and Customs Enforcement agents.

The president made the announcement during a speech he gave last night from the Old Post Office building in Washington, D.C., now a Trump International Hotel. In a surprise move after the speech, Trump invited reporters from the media to stand right next to him at the podium to field questions. "You_said_ eye face," he told NBC Nightly News reporter Megyn Kelly, who was covering the speech from a news van. "I'm going to do that because she has been placed on a White House list."

Although Trump reiterated his promise to expand the 11,000 immigration agents to a two-year-old deadline, he also said he would not immediately order details but said he intends to flesh out the policy with Congress. "We're going to do it," he said.

BREAKING NEWS

TRUMP: DEPORT ILLEGALS 'SO FAST YOUR HEAD WILL SPIN'

LIVE NOW: PRESIDENT ADDRESSES THE NATION

Markets sink as trade war looms

WORLDWIDE STOCKS plunged on Friday, continuing to fall month on record as trade war with China and Mexico seem imminent.

Markets will likely follow to the Nikkei benchmark on speculation that China is disrupting shipping lanes. Treasury Secretary Steven Mnuchin said Chinese imports and 35 percent of the U.S. manufacturing jobs.

"I don't mind trade wars when we're losing \$58 billion a year to China," Mr. Mnuchin said. But Chinese officials have made it no secret that they will put up a fight if the United States follows through with its threats.

Methods

Data Processing

- API: Networkx, NLTK
- Average clustering, transitivity, centrality, connectivity, etc.
- Algorithm to differentiate different types of users
- Text Classification (Bayes Classifiers, SVM, sentiment)
- Topic Modeling (Generative Models and LDA)

Sample size

- now 1,000 top news/worldnews and several hundred comment trees for each)

Model Selection

- Machine Learning: lasso regression, principal component regression, spline

Few Points to Mention

Algorithm Confounding Issues

- Probably only top news can be scrapped
(Reddit limits number of objects to be returned by Praw within 1000)
- Unpredictable size of comment data
- Data sources, sample size and methods are subject to change

Measurement Issues

- 1) unable to measure the impact from the offline world
- 2) silent participants are filtered
- 3) unable to study the interaction between different dimensions

Citations

Buntain, Cody, and Jennifer Golbeck. "Identifying social roles in reddit using network structure." In *Proceedings of the 23rd International Conference on World Wide Web*, pp. 615-620. ACM, 2014.

Horne, Benjamin D., and Sibel Adali. "The impact of crowds on news engagement: A reddit case study." *arXiv preprint arXiv:1703.10570* (2017).

Lippmann, Walter. *Public opinion*. Routledge, 2017.

Mrogers, Everett, and James Wdearing. "Agenda-setting research: Where has it been, where is it going?." *Annals of the International Communication Association* 11, no. 1 (1988): 555-594.

Reis, Julio, Fabricio Benevenuto, P. Vaz de Melo, Raquel Prates, Haewoon Kwak, and Jisun An. "Breaking the news: First impressions matter on online news." In *Proceedings of the 9th International AAAI Conference on Web-Blogs and Social Media*. 2015.

Zaman, Tauhid, Emily B. Fox, and Eric T. Bradlow. "A Bayesian approach for predicting the popularity of tweets." *The Annals of Applied Statistics* 8, no. 3 (2014): 1583-1611.