CS 6474/CS 4803 Social

Computing: Credibility

Munmun De Choudhury

munmund@gatech.edu

Week 10 | October 23, 2017

Midterm Presentations/Milestone Reports

- 9 groups in all
- Presentations: Oct 30
- Reports due: Nov 1
- Structure of presentations same as the reports
- Each presentation will be about 7 minutes including the talk and Q&A
 - 5 minutes of talk
 - 2 minutes of Q&A

Assignment II Discussion

Reddit Apologizes For Speculating About Boston Marathon Suspects

 The Huffington Post | By Katherine Bindley
 Y
 L

 Posted: 04/22/2013 5:18 pm EDT
 Updated: 04/23/2013 7:31 pm EDT
 Updated: 04/23/2013 7:31 pm EDT



Speculation on social media last week over who was responsible for the **bombing at** the Boston Marathon produced its own set of innocent victims: the falsely accused.

Reddit -- which was fiercely criticized for its "Findthebostonbombers" thread that called out specific people standing near the scene as suspects -- has now issued a public apology to those whose names were dragged through the mud.

"We all need to look at what happened and make sure that in the future we do everything we can to help and not hinder crisis situations," the statement reads.

Sponsored Links



Odd Trick Fights Diabetes "Unique" Proven Method To Control Blood Sugar In 3 Weeks. Watch Video. Smart-Consumer-Lifestyles.com

Know Where You Stand

 The message, which was published in a blog post on Monday, continues by mentioning the serious ramifications of the rampant speculation that occured on the site.

https://www.huffingtonpost.com/2013/04/22/reddit-boston-marathon-apology-suspects_n_3133472.html

Information Credibility on Twitter

Summary (1)

- The paper investigates use of machine learning methods to assess credibility of tweets distinguish news from rumor or fake information.
- Features used: message, user, topic, propagation.
- Findings:
 - User based features perform better than message features
 - Presence of URL in a post was the feature with highest information gain followed by estimate of negative sentiment and question-centric nature of posts.
- Human assessment to create ground truth for the algorithm.
 86% accuracy.
- Limitations:
 - Only focused on tweets belonging to bursty topics.

Summary (2)

- Event identification from Twitter Monitor
- Separate news from personal opinion (Mechanical Turk)

Credibility assessment of news (Mechanical Turk)

Distinguishing credibility levels from a set of tweets

Guidelines

Users of Twitter post short messages, each up to 140 characters, commonly known as tweets.

In this task you will need to indicate a level of credibility for the topic behind these short messages in Twitter

- We provide credibility levels: "almost certainly true", "likely to be false", "almost certainly false", and "I can't decide".
- For each group, we provide a short descriptive sentence that help you understand the topic behind the tweets. We provide also the date of the group of tweets

Examples:

News

- \$1.20 trillion deficit for 2010 confirmed.
- Vimeo, an application, is now available on the iPad.
 Spain wins the 2010 FIFA world cup in extra time

Rumors

- Hurricane in the south of Chile.
- Microsoft releases Office 2012
- Justin Bieber lyrics auctioned off for \$12 million

Item

Summary sentence: "underwood carrie

Date: Sat Jul 10 2010

Sample of messages/tweets ordered by timeline:

- · @istruckd_Annie I like all type of music from india arie , wale , kanye , carrie underwood. I like erbody :) Check this out: Carrie Underwood Wedding Takas Her Off Ne Market http://www.notsorealmess.com/carine underwood wedding Takas Her Off Ne Market http://www.notsorealmess.com/carine underwood wedding/ [1. alexis cohen] [2. dorell wright] [3. carrie underwood wedding] [4. las tablas pnama] [5. stephen colletti] congrats to my beautiful fined brittany and lovely hubby ryan on their wedding.oh and of course carrie underwood and mike fishers
- wedding!
- gonna need alot of \$ RT @sportschickblog carrie underwood married mike fisher today, not @shill910 ... i Carrie Underwood wedding (
- #np carrie underwood-temporary home
- bagoesss itu ⊤.⊤ hiksss RŤ @aayyyuuuu mandi aaahh...... ≠nowplaying I told you so carrie underwood.... gak bosen2 aku
- dengerinnyaaaa Babs Says: Carrie Underwood and Mike Fisher Wed! <u>http://www.babblewood.com/2010/07/carrie-underwood-and-mike-fisher-wed/</u>
- carrie underwood got married...i have no reason to live...
- New pix from LAX of Carrie Underwood & Mike Fisher leaving for their honeymoon! http://carrie-underwood.love.com /photos?photodeeplinkNum=0

Please classify these messages as:

Almost certainly true

- Likely to be false
- Almost certainly false
- I can't decide

Please, explain in only one sentence what made you decide (we need this to validate your HIT):

Summary (3)

	Min	Max	Mean	StdDev
AVG REG AGE	1	1326	346	156
AVG STAT CNT	173	53841	6771	6627
AVG CNT FOLLOWERS	5	9425	842	946
AVG CNT FRIENDS	0	1430	479	332
FR HAS URL	0	1	0.616	0.221
AVG SENT SCORE	-2	1.75	-0.038	0.656
FR SENT POS	0	1	0.312	0.317
FR SENT NEG	0	1	0.307	0.347
CNT DIST SHORT URLS	0	4031	121	419
SHR MOST FREQ AU	0	1	0.161	0.238
FR TW USER MENTION	0	1	0.225	0.214
FR TW QUEST MARK	0	1	0.091	0.146
FR EMOT SMILE	0	0.25	0.012	0.028
FR PRON FIRST	0	1	0.176	0.211
MAX LEV SIZE	0	632	46	114

Class	TP Rate	FP Rate	Prec.	Recall	F_1
A ("true")	0.825	0.108	0.874	0.825	0.849
B ("false")	0.892	0.175	0.849	0.892	0.87
W. Avg.	0.860	0.143	0.861	0.860	0.86



Other examples

Controversy and Sentiment in Online News

Yelena Mejova*, Amy X. Zhang[‡], Nicholas Diakopoulos[†], Carlos Castillo^{*} *Qatar Computing Research Institute, [†]University of Maryland, [‡]MIT CSAIL ymejova@qf.org.qa, axz@mit.edu, nicholas.diakopoulos@gmail.com, chato@acm.org

ABSTRACT

How do news sources tackle controversial issues? In this work, we take a data-driven approach to understand how controversy interplays with emotional expression and biased language in the news. We begin by introducing a new dataset of controversial and non-controversial terms collected using crowdsourcing. Then, focusing on 15 major U.S. news outlets, we compare millions of articles discussing controversial and non-controversial issues over a span of 7 months. We find that in general, when it comes to controversial issues, the use of negative affect and biased language is prevalent, while the use of strong emotion is tempered. We also observe many differences across news sources. Using these findings, we show that we can indicate to what extent an issue is controversial, by comparing it with other issues in terms of how they are portrayed across different media.

Our contribution. In this work, we quantify the use of emotional and biased language when presenting controversial issues in the news. We begin by building a list of controversial and noncontroversial terms in current news in the U.S. using crowdsourcing techniques. Then, we perform a large-scale analysis of millions of news articles from 15 U.S.-based news sources. We focus on the expression of sentiment using a series of lexical resources containing words conveying positive and negative emotions; this automatic analysis helps reduce the inherent subjectivity of traditional content analysis methods.

We demonstrate that controversial issues in news can be characterized by the use of fewer positive words and a greater presence of negative words. This finding is consistent across different media sources and confirmed with 4 different sentiment lexicons. Interestingly, we find that the use of highly emotional terms (as opposed to mild ones) is *less* likely in the context of controversial topics, suggesting a self-moderation on the part of the news sources.

The paper does not exploit the wealth of information embedded in the network structure of a user. This can be very useful for credibility purposes. Discuss some ways you'd utilize this information.

The authors stop short of proposing any sort of user-specific model (although they use userfeatures): beyond attributes of the information and the information sharing agent, it is necessary to model end user perceptions.

For instance, some users may value sentiment more than social connections, which may in turn impact how they assess information credibility. Are non-experts (e.g., Turkers) the right people to assess credibility?

A need for "fact checking systems" that operate outside of the social media ecosystem. But these systems are difficult to build and use. Why? Tweeting is Believing? Understanding Microblog Credibility Perceptions

Summary

- The article presents results of a survey study that gauged users' perceptions of tweet credibility.
- 26 features identified to impact credibility assessment based on a pilot sample of five individuals.
- Two controlled experiments to measure the impact of several tweet features (message topic, user name, and user image) on perceptions of message and author credibility
 - Participants recruited from inside Microsoft and through a message board of CMU
- Total of 256 completed surveys, 101 from the corporate group and 155 from the alumni group.
- Main findings:
 - Users are poor judges of tweet credibility based on content alone; reliance on cues such as username and presence/absence of profile pictures.

Other examples

Newsworthiness and Network Gatekeeping on Twitter: The Role of Social Deviance

Nicholas Diakopoulos

Columbia University, School of Journalism nicholas.diakopoulos@gmail.com

Abstract

Publishers of news information are keen to amplify the reach of their content by making it as re-sharable as possible on social media. In this work we study the relationship between the concept of social deviance and the re-sharing of news headlines by network gatekeepers on Twitter. Do network gatekeepers have the same predilection for selecting socially deviant news items as professionals? Through a study of 8,000 news items across 8 major news outlets in the U.S. we predominately find that network gatekeepers re-share news items more often when they reference socially deviant events. At the same time we find and discuss exceptions for two outlets, suggesting a more complex picture where newsworthiness for networked gatekeepers may be moderated by other effects such as topicality or varying motivations and relationships with their audience

Arkaitz Zubiaga

Dublin Institute of Tech., Applied Intelligence Research Centre arkaitz@zubiaga.org

network gatekeepers. Previous literature has found that events with high deviance were more likely to gain coverage in the main stream media (MSM) (P. Shoemaker, Danielian, and Brendlinger 1991), "the role of news media is not to mirror the world as it is, but rather to spotlight and draw public attention to problems and situations that need solutions and repair," (Pamela Shoemaker 2006). Shoemaker's theory posits that MSM will select for and favor socially deviant stories and events.

But do network gatekeepers share a similar predilection for selecting socially deviant news items? Is social deviance a professionally imbued newsworthiness criterion, or something that more generally explains interest, attention, and sharing of news? Networked gatekeeping theory explores the idea that every user on Twitter is a gatekeeper, with the discretion to share or not One limitation of the work is that their current recruitment method does not include certain

demographics that consume tweets, like teenagers or adults without a college degree; education may matter

The paper focused on a rather well-educated and specialized group of participants, and that it failed to contrast results of this population and a more general population Class Reading – Significance of assessing credibility of anti-vax information Both papers focus on assessing credibility of news. Would same observations apply to judging credibility of non-real time information? E.g., health myths

Class Exercise I

Credibility is, after all, a domaindependent attribute. What additional new feature would you consider, in addition to the ones raised in the papers, that could be useful for this purpose? How would you factor in end users' bias in perception of credibility? (Take the example of the antivax health myth)

Which of these pictures of Hurricane Sandy are Real and which ones are fake?





Amazing picture of hurricane #Sandy decending in New York pic.twitter.com/3mMhCbNq



29 Oct 12

🛧 Reply 🔁 Retweet 🔺 Favorite



L- Follow

McDonalds in Virginia Beach flooded. pic.twitter.com/FZBoCydM Reply Betweet Favorite









7 Fake Hurricane Sandy Photos You're Sharing on Social Media







29 Oct 12

🛧 Reply 1 Retweet 🔺 Favorite



L- Follow

McDonalds in Virginia Beach flooded. pic.twitter.com/FZBoCydM Reply B Retweet * Favorite









Class Exercise II

The papers examined and studied credibility on Twitter – a primarily text based content system. How would the different cues change if we look at the host of new multimedia sharing social apps (e.g., Instagram) and want to assess credibility of such content?

Modulating Video Credibility via Visualization of Quality Evaluations

Nicholas Diakopoulos School of Communication and Information Rutgers University 4 Huntington St., New Brunswick, NJ, USA diakop@rutgers.edu

Irfan Essa School of Interactive Computing Georgia Institute of Technology 75 5th St. NW, Atlanta, GA, USA irfan@cc.gatech.edu

ABSTRACT

In this work we develop and evaluate a method for the syndication and visualization of aggregate quality evaluations of informational video. We enable the sharing of knowledge between motivated media watchdogs and a wider population of casual users. We do this by developing simple visual cues which indicate aggregated activity levels and polarity of quality evaluations (i.e. positive / negative) which are presented in-line with videos as they play. In an experiment we show the potential of these visuals to engender constructive changes to the credibility of informational video under some circumstances. We discuss the limitations, and future work associated with this approach toward video credibility modulation. information quality by combing through the media and engaging in fact-checking and re-contextualization of news and other media reports. For high profile video events such as the State of the Union address given by the president of the U.S., there is a considerable demand for this type of watchdogging activity. For instance, recent coverage by news outlets like PBS included annotated transcripts and video snippets showing analysis from experts and journalists¹. One of the major issues with such analytic presentations as are found on Politifact, Factcheck, and PBS is that, especially for video, the analysis is divorced from the video itself, making the multimedia context difficult to understand in relationship with the textual analysis.

While most methods of watchdogging are labor intensive, another

