

Investigating Bias in Traditional Media through Social Media

Arjumand Younus^{*†◊}, M. Atif Qureshi^{*†◊}, Suneel Kumar Kingrani[†], Muhammad Saeed[†],
Nasir Touheed[†], Colm O’Riordan^{*}, Pasi Gabriella[◊]

^{*}Computational Intelligence Research Group, Information Technology, National University of Ireland,
Galway, Ireland

[◊]Information Retrieval Lab, Informatics, Systems and Communication, University of Milan Bicocca,
Milan, Italy

[†]Web Science Research Group, Faculty of Computer Science, Institute of Business Administration,
Karachi, Pakistan

ayounus, maqureshi@iba.edu.pk, suneel.kingrani@gmail.com, saeed@uok.edu.pk,
ntouheed@iba.edu.pk, colm.oriordan@nuigalway.ie, pasi@disco.unimib.it

ABSTRACT

It is often the case that traditional media provide coverage of a news event on the basis of journalists’ viewpoints - a problem termed in the literature as media bias. On the other hand social media have given birth to an alternative paradigm of journalism known as “citizen journalism”. We take advantage of citizen journalism to detect the bias in traditional media and propose a simple model for empirical measurement of media bias.

Categories and Subject Descriptors

J.4 [Social and Behavioral Sciences]: Sociology

Keywords

media bias, social media, named entities, topic models

1. INTRODUCTION

With the emergence of Web 2.0 there has been a tremendous increase in the usage of social Web applications which have considerably changed the nature of how media operate. This has given birth to an alternative paradigm of social media for journalism on the Web. This paradigm has caused people to shift from traditional media (newspapers, television etc.) to social media sites like Twitter to find news [1]. In fact social media have given birth to the concept of citizen journalism with ordinary citizens now playing an active role in news dissemination and discussion. It has been observed that mass media remain largely controversial in nature due to being controlled by “the elite few”. On the other hand social media provide a forum for the masses to express their concerns and opinions. A few examples are shown in Table 1.

Despite the hype surrounding social media as an alternative paradigm for journalism there has not been much investigation into how social media differ from traditional media in news reporting. Our hypothesis is based on the

notion that the alternative journalism paradigm of social media can serve as a test bed for the measurement of bias in the traditional media platforms. Using publicly available data on media sources, we conducted a study of bias in traditional media. We also propose a simple method (built on top of topic models) to measure the bias in traditional media through the use of social media.

2. QUANTITATIVE ANALYSIS OF COVERAGE FOR A MAJOR EVENT

How do traditional media outlets and social media differ in the coverage of an event? To answer this question we conducted a study of the coverage patterns of the two sources (i.e., NYTimes articles and tweets) during the Egyptian uprising in January, 2011.

For our experiments, data were gathered using the TREC 2011 microblog track and we gathered all tweets from 23rd January, 2011 to 28th January, 2011. This data significantly covers the time period of the Egyptian revolution which was an event heavily tweeted by the masses and social media activists alike. We simultaneously collected all the New York Times articles regarding the Egyptian revolution for

Date	Tweet
26th Jan	What is wrong with the TV channels?! No news about #egypt!!! #Jan25 #Arab #Media
27th Jan	A call to everyone in Egypt: Boycott all national governmental newspapers (Al Ahram, Al Akhbar, Al Gomhoreya). Spread it. #Jan25
28th Jan	RT @BrianKeene: disparity in US media Egypt coverage: CNN=peaceful protests. FOX=violent protests. MSNBC=Charlie Sheen banged porn star.

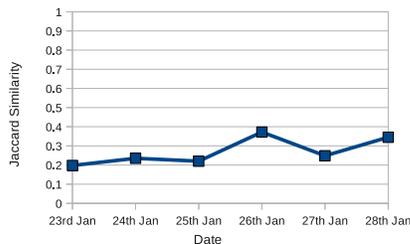
Table 1: Tweets expressing anger against media during Egypt uprising

Date	Extra Entities	Entities Missed
23rd Jan.	Muslim Brotherhood, Al-Qaeeda, Iran, Osama bin Laden	Clinton, Montreal, Canada
24th Jan.	Hamas	Jordan, Mubarak, France, Sarkozy
25th Jan.	Al Nahda, Jeffrey D. Feltman	Ben Wedeman, ElBaradei, Wikileaks, Ayman Nour
26th Jan.	Jeffrey D. Feltman, Moktada al-Sadr	Alexandria, Suez, Bradley Manning, Saleh, Ben Wedeman
27th Jan.	Al-Qaeeda, Margaret Scobey, Benjamin Netanyahu	Wikileaks, Khaled Said, Libya, Tahrir Square
28th Jan.	Margaret Scobey, Safwat el-Sharif, Ayman Nour	Julian Assange, Nour DSL, Anonymous, Gamal Mubarak

Table 2: Details of missing and extra entities between Traditional and Social Media

the same period. The time period was chosen on account of the high level of media bias exhibited at the start of the Egyptian revolution. Figure 1 shows the jaccard similarity scores for entity sets in the two media (social and traditional). As evident from the figure, the bias is quite high in terms of entity coverage and it remains below the 0.5 threshold. Table 2 provides a list of some significant extra and missing entities for the time period of our analysis.

Figure 1: Jaccard Similarity between Traditional and Social Media



3. MEASURING BIAS IN TRADITIONAL MEDIA FOR DAY-TO-DAY EVENTS

We propose a simple media bias measurement model for day-to-day news items built on top of topic models. We gather the set of article topics (from online news sources) and tweet topics (from tweets of well-known social media activists). This is done by applying the standard latent dirichlet allocation (LDA) topic model[2] to news articles and Twitter-LDA[3] to tweets. The topics extracted are then compared against each other so as to find similar topics (using an empirically set JS-divergence threshold). Similar topics found from both the datasets are then ranked as follows.

For the topics t found from news articles we perform the ranking using the metric $newsarticle_{rank}$. Here, A is the total number of articles for a particular day while $\Pr(t | a)$ is obtained from LDA.

$$newsarticle_{rank} = \frac{\sum_{a \in A} \Pr(t | a)}{|A| * js - divergence}$$

For tweets ranking is performed by simply counting number of tweets assigned to a topic.¹

We collected all tweets of Pakistani citizen journalists (social media activists) for the month of November, 2011 along

¹Note that this ranking is performed for the similar topics of news and tweets.

Date	Traditional Media	Social Media
14th Nov.	3	1
15th Nov.	4	2
19th Nov.	3	0
22nd Nov.	2	0
23rd Nov.	3	1
27th Nov.	5	2

Table 3: Ranking Difference for News Items in Traditional and Social Media

with simultaneous collection of news data from five local news sources namely Express Tribune, Dawn News, The Nation, Daily Times and The News. A manual coder prepared a manually ranked list of news topics (obtained through standard LDA) which was then compared against the lists obtained from Twitter and traditional news sources. The Spearman footrule values for both of them are given in Table 3². Traditional news sources have a wider disparity in the ranks and hence there is strong presence of media bias in the traditional media.

4. FUTURE WORK

While our model detects bias in traditional media outlets we hope that a few more considerations can improve the model to reflect various aspects of the media bias process. One significant aspect we intend to focus on as future work is sentiment analysis of social media sources.

5. REFERENCES

- [1] J. An, M. Cha, K. Gunmadi, and J. Crowcroft. Media landscape in Twitter: A world of new conventions and political diversity. In *Proceedings of 5th International AAAI Conference on Weblogs and Social Media*, pages 18–25, 2011.
- [2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022, Mar. 2003.
- [3] W. X. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li. Comparing twitter and traditional media using topic models. In *Proceedings of the 33rd European conference on Advances in information retrieval, ECIR'11*, pages 338–349, Berlin, Heidelberg, 2011. Springer-Verlag.

²The table shows a snapshot of the results.