MediaRank: Computational Ranking of Online News Sources

Junting Ye
Stony Brook University
Stony Brook, NY
juyye@cs.stonybrook.edu

Steven Skiena
Stony Brook University
Stony Brook, NY
skiena@cs.stonybrook.edu

ABSTRACT

In the recent political climate, the topic of news quality has drawn attention both from the public and the academic communities. The growing distrust of traditional news media makes it harder to find a common base of accepted truth. In this work, we design and build *MediaRank* (www.media-rank.com), a fully automated system to rank over 50,000 online news sources around the world. *MediaRank* collects and analyzes one million news webpages and two million related tweets everyday. We base our algorithmic analysis on four properties journalists have established to be associated with reporting quality: peer reputation, reporting bias/breadth, bottomline financial pressure, and popularity.

Our major contributions of this paper include: *(i)* Open, interpretable quality rankings for over 50,000 of the world's major news sources. Our rankings are validated against 35 published news rankings, including French, German, Russian, and Spanish language sources. *MediaRank* scores correlate positively with 34 of 35 of these expert rankings. *(ii)* New computational methods for measuring influence and bottomline pressure. To the best of our knowledge, we are the first to study the large-scale news reporting citation graph in-depth. We also propose new ways to measure the aggressiveness of advertisements and identify social bots, establishing a connection between both types of bad behavior. *(iii)* Analyzing the effect of media source bias and significance. We prove that news sources cite others despite different political views in accord with quality measures. However, in four English-speaking countries (US, UK, Canada, and Australia), the highest ranking sources all disproportionately favor left-wing parties, even when the majority of news sources exhibited conservative slants.

ACM Reference Format:
Junting Ye and Steven Skiena. 2019. MediaRank: Computational Ranking of Online News Sources. In *The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '19), August 4–8, 2019, Anchorage, AK, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3292500.3330751

1 INTRODUCTION

A common base of accepted truth is perhaps the most important foundations of democracy, yet this has come under assault in our era of fake news and the widespread distrust of traditional media. Considerable work has been devoted to developing NLP-based methods to detect unreliable news articles [17, 20], as well as independent third-party fact checking services like *PolitiFact* and *Snopes*, but validity checking on the article level is too brittle and slow relative to the demands of the news cycle.

We believe that the proper level to assess news quality is at the source level, through aggregate analysis of their coverage, content, and reputation. While the professional journalists offer accurate annotations on evaluating the quality of news sources (e.g. *NewsGuard*), it is difficult and expensive for them to achieve high coverage due to the sheer amount of information generated everyday.

Towards this end, we have developed *MediaRank* (www.media-rank.com), a fully automated system to rank over fifty thousand online news sources around the world. We collect and analyze about one million new webpages and two million related Tweets everyday. This longitudinal dataset represents a substantial academic resource for analyzing news media and information flow around the world.

Ranking online news sources proves a challenging task. A straightforward approach one might use is traditional website ranking algorithms, e.g. *PageRank* [16]. But as we will show in Section 4.2, this does not prove an effective approach because of "sponsored articles" and other uninformative hyperlinks that dominate news pages. Instead, multiple metrics must be considered to assess the media quality. According to surveys of top U.S. journalists conducted by *Pew Research Center*, political balance journalism, quality of coverage (e.g. depth and context) and bottomline pressure are among the key factors influencing the quality of news sources [19].

With this domain wisdom in mind, we propose the following four properties to assess the quality of news sources, and develop novel algorithmic methods to evaluate them:

- *Peer Reputation*: Reliable news sources are trusted by other reliable news sources. Reporting citations are common in online news articles. We argue that news sources receive more

Rankings	Media Rank	Nuzzel Rank[1]	News Guard[2]	Feed Spot[3]	AllYou CanRead[4]
Public	✔	✘	✔	✔	✔
Multi-Topics	✔	✔	✔	✔	✔
Multi-Lang	✔	✔	✘	✔	✔
Multi-Nation	✔	✔	✘	✔	✔
>50K Sources	✔	✔	✘	–	✘
Interpretable	✔	–	✔	✘	✘
Algorithmic	✔	–	✘	✘	✘

Table 1: Comparisons of *MediaRank* against other news ranking systems: *NuzzelRank, NewsGuard, FeedSpot,* and *AllYouCanRead.* Blank entries reflect lack of reliable information concerning methodology and coverage.

[1]https://nuzzel.com/rank
[2]https://www.newsguardtech.com/
[3]https://www.feedspot.com/
[4]https://www.allyoucanread.com/

KDD '19, August 4–8, 2019, Anchorage, AK, USA

ACM ISBN 978-1-4503-6201-6/19/08…$15.00
https://doi.org/10.1145/3292500.3330751

News *MediaRank* Favors					News *NuzzelRank* Favors				
News	Δ	$\widetilde{MR}$	NR	MR	News	Δ	NR	$\widetilde{MR}$	MR
variety.com	76	16	92	17	mediamatters.org	-52	42	94	4290
nature.com	62	14	76	15	qz.com	-47	37	84	1005
sciencemag.org	56	37	93	51	gizmodo.com	-46	13	59	121
rollingstone.com	46	45	91	65	fastcompany.com	-44	26	70	211
independent.co.uk	45	24	69	29	thedailybeast.com	-44	25	69	204
telegraph.co.uk	42	20	62	21	entrepreneur.com	-40	35	75	316
apnews.com	39	21	60	25	propublica.org	-40	38	78	381
usatoday.com	38	10	48	10	venturebeat.com	-38	59	97	6521
scmp.com	37	50	87	87	zdnet.com	-36	30	66	177
espn.com	37	8	45	8	bostonglobe.com	-35	19	54	91

Table 2: Contrasting the top 10 news sources with biggest ranking gaps between *MediaRank* (MR) and *NuzzelRank* (NR). $\widetilde{MR}$ is induced *MediaRank* value among the 97 available news from NR.

citations from good places have higher reputation. Therefore, we use *PageRank* scores on reporting citation graph to evaluate the importance of news sources. This metric proves to be particularly effective for large-scale news sources.

- *Reporting Bias and Breadth*: Reliable news sources strive to be politically unbiased in their search for truth. Further, they strive to cover the full breadth of important news rather than repeated coverage of narrow domains. We measure reporting bias by the sentiment differences towards a large universe of people associated with left- and right-wing parties. The magnitude of sentiment bias can be accurately quantified through longitudinal analysis over a large news corpus. Breadth of reporting is estimated by the count of unique celebrities' names mentioned in their articles.
- *Bottomline Pressure*: The business environment for news venues has become increasingly challenging, with most sources facing considerable financial pressure to attract and monetize readers. But bottomline pressure is regarded by journalists as the biggest concern affecting news quality [19]. We propose two new metrics to assess integrity under financial pressure: *(i)* the use of social network bots hired to boost user traffic, and *(ii)* the number and placement of ads shown on news pages to gain revenue.
- *Popularity*: More reliable news sources are recognized as such by readers and other news sources. Social media and content analysis links and Alexa rank scores[5] reflect the popularity among news readers and sources. We demonstrate that popularity correlates strongly with peer reputation but is independent of bias.

MediaRank combines scores from the signals described above to compute a quality score for over 50,000 sources. Table 1 compares our methodology *MediaRank* to other new ranking systems, establishing us as the only large-scale, international, algorithmic news ranking system with publicly released rankings for evaluation and analysis. Table 2 compares our source rankings to *NuzzelRank*, perhaps the most comparable system, but one that releases only the relative rankings of its top 99 sources. Although there is general agreement (Spearman correlation 0.52) the differences are revealing when we identify the most disparate rankings among their sources. We strongly prefer the sources of record in science (*Nature* and *Science*) and entertainment (*Variety*, *Rolling Stone*, and *ESPN*), and professional news sources (the *Associated Press*, *Independent*, and *Telegraph*), while *NuzzelRank* favors blog-oriented sources like *VentureBeat*, *QZ*, and *Media Matters*.

The major contributions in this paper are:

- *Open, interpretable quality rankings for the world's major news sources* – We provide detailed computational analysis for over 50,000 news sources from around the world. We evaluate our rankings against 35 published news rankings, including French, German, Russian, and Spanish language sources. *MediaRank* scores correlate positively with 34 of 35 of these expert rankings, achieving a mean Spearman coefficient of 0.58. We concur with 24 of these expert rankings at above a 0.05-significance level, with a mean coefficient 0.69. Each source ranking score can be interpreted by six intuitive metrics regarding reputation, popularity, quality of coverage and bottomline pressure. We will make this analysis fully available to the research community and general public at www.media-rank.com.
- *New computational methods for measuring influence and bottomline pressure/social bots* – To the best of our knowledge, we are the first to study the large-scale news reporting citation graph in-depth. We are also the first to study computational ways to measure the bottomline pressure among news sources. Observing online news make most of their revenue from user traffic and online advertisement, we propose methods to detect social bots that promotes website traffic and to track the volume and aggressiveness of advertisements on news webpages. These metrics present interesting views into the business of the media world, and new tools for analyzing other websites and social media properties.
- *Media bias and significance* – We have performed extensive experiments using our signal metrics to quantify properties of media sources, with interesting results. In particular, we prove that news sources cite others despite different political views (Figure 2) in accord with quality measures. We also were surprised to learn that neutral sources were *not* those most highly ranked by other metrics. Indeed, in four English-speaking countries (US, UK, Canada, and Australia), the highest ranking sources all disproportionately favor left-wing parties, even when the majority of news sources exhibited conservative slants (Figure 3).

2 RELATED WORK

The problem of news source ranking has been attracting growing attention from academic and industrial researchers. Corso et. al. studied the problem of simultaneously ranking news sources and its stream of news articles [3]. They proposed a graph formulation where nodes are news sources and articles. The edges reflect relations between sources and articles, and content similarity between articles. A time-aware label propagation algorithm is proposed to assign weights to nodes in this graph. Mao and Chen suggested a similar approach to simultaneously rank news sources, topics and articles, assuming that trust-worthy news sources publish high-quality articles concerning important news topics [14]. Hu et. al. analyzed the visual layout information of news homepages to exploit the mutually reinforcing relationship between news articles

[5]https://www.alexa.com/siteinfo

and news sources [9]. These methods are dependent on computationally expensive models over articles, like label propagation. Therefore they are limited to small news corpora, and not appropriate for datasets with hundreds of millions of articles like ours.

NuzzelRank is a news recommendation system which also generates rankings of news sources. They claim their scores are computed by combining the reading behavior of their users, the engagement and authority of news sources and signals from news reliability initiatives such as the *Trust Project*[6] and *NewsGuard*[7]. We identified their top 99 ranked news sources (all that they made available to the public as of Oct. 23, 2018) for comparison with *MediaRank*.

Online misinformation is now drawing increased attention from the research community [17, 20, 21, 24, 28]. Zhang et.al. define credibility indicators in news articles for manual annotation, including eight content (e.g. title representativeness, quotes from outside experts, etc.) and eight context indicators (e.g. originality, representative citations, etc.) [30]. Linguistic models achieve limited performance in detecting fake news, especially the ones aim to deceive readers [17]. A hybrid model combining news text, received responses and the source users promoting them is proposed by [21]. Online misinformation spreads quickly on social media platforms, due to the convenience of message sharing [23]. Algorithms designed to take down social bots who publish or share misinformation or other content automatically include [1, 2, 13, 28].

Substantial efforts have been made to analyze and rank individual news articles by information retrieval community [11, 26]. Kiritoshi and Ma rank news articles by estimating the relatedness, diversity, polarity and detailedness of its named entities [10]. Tatar et. al. uses user comments to predict the popularity of news articles [25]. Godbole et. al. propose efficient algorithms for large-scale sentiment analysis of online news and social media [7]. Kulkarni et. al. design a multi-view attention model to classify the political ideology of news articles [12].

3 *MEDIARANK* OVERVIEW

The lack of valid ground-truth labels makes news ranking a challenging task. In this work, we design effective and interpretable component signals from different perspectives regarding news quality. This makes it easy to explain why one source is better than the other. Considering the sheer amount of news data everyday, each signal metric we use has been designed so be scalable for large-scale data analysis.

MediaRank is a large system, with 1 master server and 100 dedicated slave servers processing the world's news. It is organized in following four major components:

(1) *News source discovery*: two strategies are employed to identify new sources: i). new URLs appear on Google News, and ii) new URLs appear in Tweets returned by *Twitter API* when searching with keyword "news". Between Sep. 24, 2017 and Oct. 30, 2018, 50,834 unique news sources are discovered in this way, with 87% of our tracked sources being from *Google News*. The remaining 13% sources identified from Twitter prove less well-known, but sometimes go viral in social media.

[6]https://thetrustproject.org

[7]https://www.newsguardtech.com

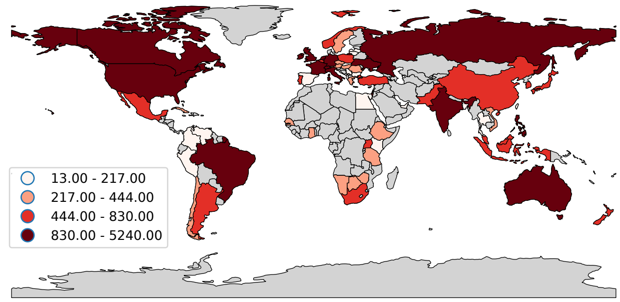

Figure 1: *MediaRank* tracks 50,696 news sources from 68 countries. Colors represent the number of sources per country. 5,240 sources are from United States. Countries with zero tracked news sources are marked in grey.

(2) *Collecting news webpages and related tweets*: We use *Newspaper3k*[8] to collect and parse news webpages from discovered domains. We also extract URLs from collected tweets to see whether it is tracked in *MediaRank*. If yes, we further query its user profile data from Twitter and keep them for analysis. On average, *MediaRank* collects about one million raw HTMLs and two million news related tweets each day. A cluster of 20 machines performs data collection and cleaning.

(3) *Analysis and News Ranking*: Multiple signals have been shown to be correlated with the quality of news sources, including reputation among peers, the degree of political bias, and popularity among readers. We devote a cluster of 80 machines to computation-intensive analysis, including named entity recognition, sentiment analysis, social bots classification, and duplicate article detection.

(4) *Visualization and API*: Our goal is to make *MediaRank* an important data source to support external research efforts in journalism and the social sciences as well as computer science. We are designing APIs to provide online service, notifying Web users whether the news they consume are from low quality sources.

Figure 1 shows the national distribution of tracked sources, using meta data from *Google News*. We observe that most sources are from western countries, with limited data from Africa and Middle East. Fully five thousand sources are from United States. Italy, Russia, Canada and U.K. are next four countries in terms of source frequency. Only 36% of our sources publish in English. Multi-language sources like *BBC* are labeled as per which language is used in the most articles. Sources with multiple topics, like the *New York Times*, are labeled as "General".

4 NEWS CITATIONS

Just as academic papers cite other papers, online news articles often acknowledge their peers' work as information sources. We argue that such citations can generally be viewed as endorsements among journalism peers. To the best of our knowledge, we are the first to generate large-scale news citation graphs for in-depth analysis and news ranking.

[8]https://github.com/codelucas/newspaper

MediaRank tier	doc/day	C_s^{out}/doc	C_o^{out}/doc	C_o^{in}/doc
[1,500)	47.0	2.8	1.0	201.0
[500, 2K)	17.3	2.0	0.9	60.3
[2K, 5K)	9.0	1.7	0.9	18.4
[5K, 10K)	5.1	1.6	0.9	8.8
[10K, 20K)	3.3	1.5	0.8	1.8
[20K, 50K)	2.3	1.3	0.8	0.2

Table 3: Higher ranking news sources *(i)* publish more articles each day, *(ii)* have more citations to both articles of their own and other sources, and *(iii)* receive more citations from others. Sources are grouped into six tiers based on their *MediaRank* values. C_s^{out}: count of self-citations, C_o^{out}: citations to other sources, C_o^{in}: citations from others.

In this section, we analyze citation behavioral patterns of news sources (Table 3). We also define the news citation graph, where the nodes are news sources and directed edges represent citations between source pairs.

4.1 Dataset

We analyze 23,371,264 articles collected between Sep. 24, 2017 and Feb. 16, 2018. Each article contains at least one citation inside it, for a total of 64,976,942 citations. Of these, 42,734,224 (66%) are self citations to given news source, while 22,242,718 (34%) cite different news sources.

The news citation graph is a directed graph, denoted as $G_c =< V, E, W >$ where news sources are the nodes, V. $e_{ij} \in E$ is a directed edge from node v_i to v_j ($v_i, v_j \in V$). The total number of citations from v_i to v_j defines $w_{ij} \in W$ the weight of edge e_{ij}, Our weighted source citation graph contains 50696 nodes and 1,947,189 edges after removing self-loop edges.

4.2 Citation Ranking

PageRank was famously defined as an algorithm to rank websites [16]. The key idea is that every webpage propagates their weight to their neighbors. When a page has many links from large-weight webpages, the weight of this page increases. Similarly, we argue that citations between news sources should be interpreted as endorsements among journalists. When a news source is disproportionately cited by its peers, it indicates a higher journalistic reputation.

We compare *PageRank* results on both citation graph and URL graph (where sources are connected by all URLs, instead of just inside articles.). By comparing the top 10 news from both rankings, we observed that certain sources ranked disturbingly higher in the URL graph than in citation graph. For example, "digg.com" stands 6th on URL ranking, while only 71st in citation ranking. "bna.com" is placed 8th on URL ranking vs. 1157th on citation ranking. The primary reason for such anomalies is that outside article links are often ads or "sponsored" articles, which prove much less informative than reporting citations.

We use *PageRank* values from the citation graph to quantify peer reputation, normalized to be in the range [0, 1]. The greater the reputation score is, the better the source is presumed to be.

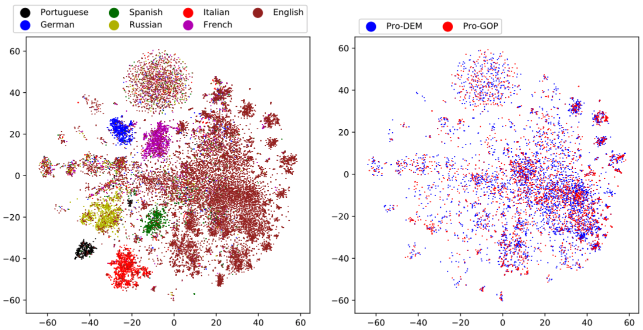

Figure 2: 2D projection of news embeddings learned from the citation graph. (Left): the colors are labeled based on news sources' languages. Strong clusters are formed by all major languages. (Right): the colors are labeled based on political news sources' sentiment towards U.S. parties. No large clusters are observed, indicating that news sources cite each other despite different political views.

4.3 Citation News Embeddings

Graph embeddings are low-dimension vector representations for nodes so that similar nodes have similar representations [18]. We are interested in how news sources align in embedding space, and what their nearest neighbors look like. We used *Node2Vec* [8] to learn news sources embeddings on news citation graph, and projected these embeddings into two dimensions for visualization purpose using *t-SNE* [27]. For visualization purposes, we used metadata from *Google News* to label sources by topics and annotated sources by political bias as explained in Section 5.

Figure 2 (left) shows the news sources distribution of top seven languages tracked in *MediaRank*. All languages form strong clusters. English proves widely used in many countries, so we see multiple smaller national sub-clusters. For the right figure, it is interesting that no large clusters are found among political news sources. This indicates that sources with different political views do cite each other, contradicting the "echo chamber" effect associated with social media platforms [5].

5 NEWS BIAS

News bias is a critical metric reflecting the quality of news sources. According to *Pew Research Center* survey of 38 countries, a median of 75% per nation say it is never acceptable for a news organization to favor one political party over others [15].

To facilitate large-scale text analysis, we employ efficient and effective algorithms to extract named entities [4] and compute sentence-level sentiment [6]. The political bias of a new source is computed by aggregating sentiments towards party members.

5.1 Datasets

We analyzed news articles collected from Sep. 24th, 2017 to Dec. 31, 2018, with 427,464 distinct celebrities are mentioned at least once. There are 77,596,029 articles containing at least one celebrity's name, totaling 614,440,328 mentions. These celebrities' English names are mapped to entities extracted from *DBpedia Data Set 3.1*[9].

[9]https://wiki.dbpedia.org/data-set-31

We identified each celebrities' political party label from DBpedia, with 2,908 unique parties are associated with 58,131 celebrities, of which 12,784 are U.S. Republicans and 11,774 Democrats. We have enriched this with Trump's cabinet (past) members[10], 115th[11] and 116th[12] class of congress members for analysis.

We identified two external resources to prove a ground-truth for news bias evaluation:

- *AllSides*[13]: 222 raw news sources with their political bias. Each source is labeled with one of following five political views by news editors: *left*, *left-center*, *center*, *right-center* and *right*, these labels are also voted by Web users. After filtering out those not tracked by *MediaRank*, 117 news sources remained. We also observed that there is often an inconsistency between the opinions of news editors and Web users. We removed the inconsistent sources and those labeled as "center", leaving 71 news sources for evaluation.
- *MediaBiasFactCheck* (*MBFC*)[14]: contains 1040 news sources labeled as "Left Bias", "Left-center Bias", "Right-center Bias" and "Right Bias". Of these, 653 are tracked in *MediaRank*. We combined "Left Bias" and "Left-center Bias" in news as "Left" and "Right-center Bias" and "Right Bias" as "Right".

5.2 Sentiment Aggregation

We now explain the details of how the sentiment of news entities and sources are computed. We consider three ways to aggregate news sentiment:

- *Article-level bias Vote (AV)*: each article has one vote towards an entity: positive, negative or neutral. The group sentiment is aggregated by counting votes from articles containing a party member.
- *Article-level bias Distribution (AD)*: similar to AV, but aggregating entity sentiment distributions instead of votes.
- *Sentence-level bias Distribution (SD)*: similar to SD, but assigning weights proportional to entity mentions instead of articles.

Formally, we assume news source $s_i =< d_1, d_2, ..., d_n >$ consists of a sequence of articles. Each article, $d_j =< g_1, g_2, .., g_m >$, consists of a sequence of sentences. Let $E_k =< e_1, e_2, .., e_u >$ denote the list of entities occurring in sentence k. Let $O(g_k)$ denote the sentiment probability distribution of sentence g_k. The distribution has three classes, positive, neutral and negative sentiments. For example, $O(g_k) = [0.1, 0.9, 0.0]$, where the entries are positive, neutral and negative sentiment scores, respectively. For each entity, its party affiliation $P(e_l), e_l \in E_k$, can be one of the 2,908 parties or none. The average sentiment distribution of party p_u from article d_j is defined:

$$O(d_j, p_u) = \frac{1}{N_d} \sum_{g_k \in d_j} \sum_{e_l \in E_k} O(g_k) * I(P(e_l) = p_u) \quad (1)$$

where N_d is the normalization term that makes $O(d_j, p_u)$ a probability distribution. $I(\cdot)$ is an indicator function, whose value is one if the condition is satisfied, otherwise 0. $O(g_k)$ is viewed as vector when under adding or multiplying operations. An article's sentiment towards a political party is the average sentiment of its sentences. $\overline{O}(d_j, p_u)$ denotes the vote of article d_j on party p_u. This one-hot vector denote whether it is a positive, neutral or negative sentiment vote. For example, $\overline{O}(d_j, p_u) = [0, 1, 0]$ is a neutral vote if the positive of $O(d_j, p_u)$ equals the negative. It takes $[1, 0, 0]$ if the positive of $O(d_j, p_u)$ is larger than the negative sentiment, otherwise $[0, 0, 1]$.

We aggregate the article-level vote as:

$$O_{av}(s_i, p_u) = \frac{1}{N_{av}} \sum_{d_j \in s_i} \overline{O}(d_j, p_u) \quad (2)$$

where N_{av} is the normalization term that makes $O_{av}(s_i, p_u)$ a probability distribution. The article-level aggregate distribution $O_{ad}(n_i, p_u)$ is defined similarly using $O(d_j, p_u)$. The sentence-level aggregate distribution is computed:

$$O_{sd}(s_i, p_u) = \frac{1}{N_{sd}} \sum_{d_j \in s_i} \sum_{g_k \in d_j} \sum_{e_l \in E_k} O(g_k) * I(P(e_l) = p_u) \quad (3)$$

Finally,the sentiment score a news source for a political party is computed:

$$B(s_i, p_u) = \frac{O^{pos}(s_i, p_u) - O^{neg}(s_i, p_u)}{O^{pos}(s_i, p_u) + O^{neg}(s_i, p_u)} \quad (4)$$

where $O^{pos}(s_i, p_u)$ and $O^{neg}(s_i, p_u)$ are the positive and negative values of sentiment distribution $O(s_i, p_u)$. $B(s_i, p_u)$ is in the range $[-1, 1]$. The absolute gap between sentiment scores of left- or right-wing parties is used to quantify source bias.

Method	EntitySet	Entity#	AllSides	MBFC
Random	–	–	0.489	0.502
Article Vote	Cabinet	27	0.507	0.509
Article Distri.	Cabinet	27	0.493	0.540
Sentence Distri.	Cabinet	27	0.541	0.526
Article Vote	Congress	564	0.701	0.540
Article Distri.	Congress	564	0.656	0.530
Sentence Distri.	Congress	564	0.666	0.557
Article Vote	All	18773	0.761	0.643
Article Distri.	All	18773	0.746	0.649
Sentence Distri.	**All**	**18773**	**0.764**	**0.683**

Table 4: Accuracies of news source bias prediction. MBFC: MediaBiasFactCheck.com.

5.3 News Bias Evaluation

To evaluate our methods for political bias detection, we used source bias labels from two organizations, *AllSides* and *MediaBiasFactCheck* (i.e. *MBFC*), as ground-truth data. Table 4 shows how our various sentiment methods perform using different groups of party-associated entities. Accuracy increases when using larger sets of party-associated entities for all aggregation methods. SD aggregation slightly outperforms other methods.

Table 5 presents the most significant left and right-leaning news sources, where the gap between democratic and republican bias > 0.05. There is excellent agreement with *MBFC* bias labels. The

[10]https://en.wikipedia.org/wiki/Cabinet_of_Donald_Trump
[11]https://en.wikipedia.org/wiki/115th_United_States_Congress
[12]https://en.wikipedia.org/wiki/116th_United_States_Congress
[13]https://www.allsides.com/media-bias/media-bias-ratings
[14]https://mediabiasfactcheck.com

News	Democratic		Republican		MBFC Label
	Bias	#(K)	Bias	#(K)	
latimes.com	+0.06	93	+0.00	264	left-center
businessinsider.com	+0.07	87	+0.00	256	left-center
theconversation.com	+0.13	5	+0.05	17	center
fortune.com	+0.10	15	+0.05	44	right-center
smh.com.au	+0.10	17	+0.03	51	left-center
usnews.com	+0.10	37	+0.05	119	left-center
vice.com	+0.04	9	-0.01	31	left-center
indiatimes.com	+0.20	3	+0.08	13	left-center
qz.com	+0.12	8	+0.06	22	left-center
miamiherald.com	+0.08	10	+0.01	21	left-center
sky.com	-0.14	7	-0.06	21	left-center*
breitbart.com	+0.03	140	+0.09	380	extreme-right
nationalreview.com	+0.02	44	+0.08	98	right
dailycaller.com	+0.02	58	+0.08	129	right
torontosun.com	-0.11	8	+0.00	18	right
eveningtimes.co.uk	-0.05	4	+0.04	10	–
clarionledger.com	+0.06	13	+0.15	44	–
abc7.com	+0.02	5	+0.09	15	–
dailyecho.co.uk	-0.04	2	+0.01	8	–
nationalinterest.org	+0.05	5	+0.12	27	right-center

Table 5: Successfully discriminating the ten most significant left- and right-wing news sources by sentiment. *Note that *Sky News* is owned by *21st Century Fox*, and considered a conservative source by *Wikipedia*.

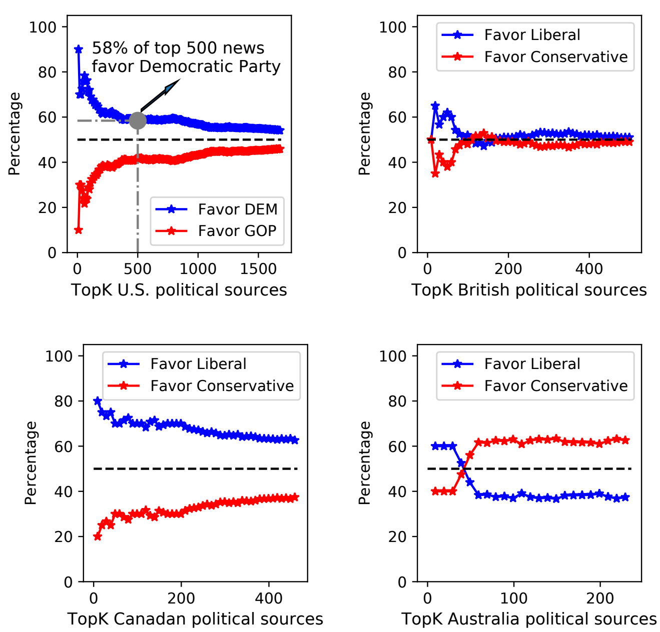

Figure 3: Bias of national news towards liberal and conservative parties in major English-speaking countries (U.S., U.K., Canada and Australia). The best news sources tend to favor liberal parties in all these countries.

outlier is *Sky News* labeled as left-center by *MBFC* but owned by *21st Century Fox*, and considered a conservative source by *Wikipedia*[15].

[15] https://en.wikipedia.org/wiki/Sky_News

6 SOCIAL BOT SCORE

Social media has become the primary vehicle for news consumption: 62% of U.S. adults received news on social media in 2016[16]. Social media outperforms television as the primary news source for younger generation (18 to 24 year old)[17]. Unfortunately, social media has also become the major outlet for distributing fake news [21], because the "echo chamber" effect makes fake news seem more trust-worthy [22]. Social bots are social media accounts controlled by computer programs. They are often used to promote public figures by following them, or to boost business by sharing related posts. It was reported that up to 15 percent of Twitter accounts are in fact bots rather than people [29]. In this section, we will elaborate on how we train a social bot classifier and further compute the social bot score of news sources.

6.1 Dataset

Twitter is one of the most popular social media platforms, and provides an API[18] enabling us to identify the user ID, tweet content, related URL, and post timestamp for millions of tweets. We used the keyword "news" in API queries to identify news-oriented tweets, and extracted all news-oriented URLs from these tweets. Between Sep. 29, 2017 and Oct. 30, 2019 (397 days), we collected 715,050,598 tweets with URLs, of which 347,164,578 (48.6%) contain URLs from tracked news sources. These Tweets are posted by 32,275,806 users, whose profiles are also collected for social bot identification.

We identified two datasets of social bot labels for training and evaluation:

- *Botometer*: this dataset is the combination of four public social bots datasets from the research community[19] [1, 2, 13, 28]. Bot labels are collected using "honeypot" (i.e. followers of accounts that post random words), or by followers bought from companies. This dataset contains 46,459 total accounts, split between 24,267 social bots and 22192 regular users.
- *Removed Accounts*: Twitter strives to remove social bot accounts[20]. We identified deleted accounts (enriched in bots) by retrieving the same user profiles twice (on Oct. 1st, 2017 and Mar. 21, 2018). Among the original user set of 1,105,536 accounts, fully 45,654 (4.1%) were not available after six months.

6.2 Social Bots Detection

We model bots detection as a supervised classification problem, using 12 features extracted from user profiles. Although follower and followee relations have proven useful in previous studies, this was not feasible on *MediaRank* scale due to Twitter API rate limits. The features we use are defined in Table 6.

The distribution of twitter account labels is highly imbalanced (only 4.1% as removed). We sampled 45,654 non-removed accounts as negatives for training. Both datasets were split 70% for training, 10% for parameter tuning and 20% for testing. As shown in Table 7, XGBoost consistently outperforms SVM classifier with RBF kernel

[16] http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/
[17] www.bbc.com/news/uk-36528256
[18] https://developer.twitter.com/en/docs.html
[19] Download: https://botometer.iuni.iu.edu/bot-repository/datasets.html
[20] https://www.theverge.com/2018/3/11/17107192/twitter-tweetdecking-spam-suspended-accounts-mass-retweeting

#	Feature	Note
1	c_{er}	Count of followers
2	c_{ee}	Count of followees
3	$r = c_{ee}/c_{er}$	Ratio of followee count over followers
4	$s = \log(\max(c_{er}, c_{ee}))$	Log of follower or followee count
5	$r * s$	Ratio times the log of follower/followee
6	v	Whether the user is verified
7	c_f	Favourites count
8	c_l	Listed count
9	c_d	The length of profile description
10	Geo	Whether geo is enabled
11	Location	Whether location is specified
12	Time zone	Whether time zone is specified
13	Default profile	Whether default profile background is changed
14	Default profile image	Whether default profile background image is changed

Table 6: Features form social bots classification model using users' profile data.

Model	*Botometer*			*RemovedAccounts*		
	Pre.	Rec.	F1	Pre.	Rec.	F1
LR	0.81	**0.85**	0.83	0.65	**0.69**	0.67
SVM	0.84	**0.85**	0.84	0.75	0.61	0.67
XGBoost	**0.88**	0.84	**0.86**	**0.79**	0.60	**0.68**

Table 7: Performance comparisons of logistic regression (LR), SVM and XGboost on two different social bot datasets.

and logistic regression model with ridge regularization on both datasets.

6.3 News Bot Scores

Therefore, we employed XGBoost models trained on two datasets to all 32 million users to get their social bot scores. Bot scores of news sources are computed by aggregating the scores for all related Twitter accounts. Sources with high bot scores likely that it hires bots to increase their visibility.

To be precise, let b_u be the bot score of Twitter user u and $T_i =< t_{i1}, t_{i2}, ..., t_{in} >$ denote the sequence of tweets with URLs directing to news source s_i. Let $U(t)$ denote the user of tweet t. Therefore, the bot score $B(s_i)$ of source s_i is defined:

$$B(s_i) = \frac{1}{|T_i|} \sum_{t \in T_i} b_{U(t)} \tag{5}$$

We combine the two models from *Botometer* and *RemovedAccounts* by using the larger of the respective scores.

7 OTHER SIGNALS

7.1 Popularity

Alexa Rank is used to estimate news sources' popularity among news readers. We collected ranking values of all sources on Sep. 23rd, 2018, using their API to collected data for the past 30 days. The average ranking values of 30 days are computed to measure its popularity.

Alexa ranks range from 1 to 1,000,000, which we divide 20 equal-range tiers. The top tier features sources with Alexa ranks between 1 to 50,000, including 6,932 (14%) of the 50K news sources tracked by *MediaRank*.

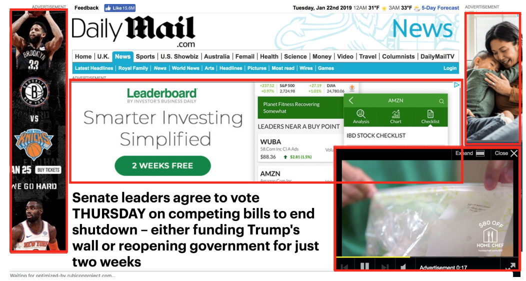

Figure 4: The *Daily Mail* is a notoriously aggressive advertiser, here with 20 digital advertisements overwhelming the news title.

7.2 Advertisement Aggressiveness

Online advertising is the major revenue stream for many news sources. Media properties under great bottom-line pressure may increase the presence of ads on their pages, reducing user experience to gain more reader clicks/impressions to survive.

To collect news advertising data, we used *Selenium*[21] to discover rendered iFrames from *Google Ads* platform in HTMLs. This was effective in terms of precisions, but less so in recall.

As an example, Figure 4 shows the first screen of a webpage from the *Daily Mail*, a popular British news source. The four observable ads here are distracting, making it hard to notice the news titles on the bottom of the page. We encountered the *Daily Mail* articles with as many as twenty ads per page, making it an example of advertising aggressiveness.

7.3 Reporting Breadth

The breadth of coverage is an important indicator of news quality, reflecting the scope, relevance, depth insight, clarity, and accuracy of reporting [19]. We use the number of unique entities to measure the breadth of news reporting. Good news sources strive to cover the full breadth of important news, rather than narrow domains with limited and repeated entity occurrence.

8 CONSENSUS SOURCE RANKING

8.1 Methodology

In our ranking model, each news source is represented by a vector of signal scores: reputation, popularity, reporting breadth, political bias, bot score and advertising aggressiveness, denoted by f_r, f_p, f_e, f_b, f_s and f_a respectively. For the four continuous signals $F = [f_r; f_p; f_e; f_b;]$ (each normalized in the range of [0, 1]), the source ranking score is defined:

$$R(s_i) = W^T \cdot F_i \cdot C_p^{I(f_s=1)+I(f_a=1)} \tag{6}$$

where $W = [w_r; w_p; w_e; w_b;]$ is the weight vector for these signals and W^T is the transpose of W. C_p is the penalizing factor to discount the weights of sources employing social bots and displaying excessive ads, measured as binary (0 or 1) features using 95 percentile values as thresholds. $I(\cdot)$ is an indicator function whose value is 1 iff the condition is satisfied. Empirically we set the $W = [1.65; -0.35; 0.05; -0.10;]$ and $C_p = 0.95$, reflecting the monotonicity of each feature.

[21] https://www.seleniumhq.org/

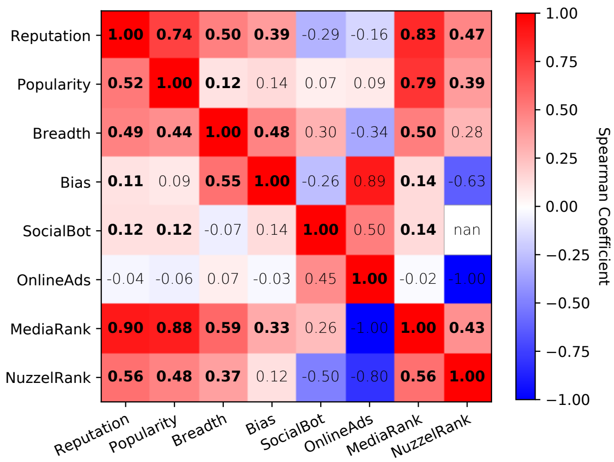

Figure 5: Spearman rank coefficients between pairs of signals, *MediaRank* and *NuzzelRank* on common sources. All coefficients with > 0.05 significance are in bold.

8.2 Evaluation

Figure 5 presents the Spearman rank correlation between signal pairs and two source rankings (*MediaRank* and *NuzzelRank*) on common sources. Because vastly more low quality news sources than high quality outlets, we use stratified sampling to compute correlations. These samples are drawn from six news tiers as sorted by *MediaRank* scores, with boundaries of rank 100, 400, 1600, 6400 and 25600. We have sampled 100 news sources from each tier. We compare the *MediaRank* rankings of the 600 sampled news to their *NuzzelRank* rankings. When comparing *NuzzelRank* to *MediaRank*, the sampled news are different, thus the coefficient matrix is not symmetric. For bot and ads scores, we use the gaps to thresholds as ranking values (news with zeros are ignored). Reputation, popularity and breadth scores highly correlate with each. Coefficients of bot, ads and *NuzzelRank* prove less significant due to smaller number of associated news sources.

In addition, we compare *MediaRank* scores with 35 expert news sources rankings (including French, German, Italian, Russian, and Spanish language sources). We also propose a ranking quality metric to quantify how good the selection of news are when comparing to *MediaRank*. Let

$$Q(S) = \sum_{s_i \in S} \frac{1}{rank_{s_i}}, rank_{s_i} \in [1, m] \tag{7}$$

$$Q_n(S) = \frac{Q(S) - Q(S_{min})}{Q(S_{max}) - Q(S_{min})} \tag{8}$$

where $rank_{s_i}$ is the *MediaRank* value of news source s_i among m sources tracked in *MediaRank*. $Q(S_{min})$ is the smallest value of $Q(\cdot)$ because news in S_{min} stand at the bottom of *MediaRank*. Similarly, $Q(S_{max})$ gets the largest value when S_{max} is at the top. Therefore, $Q_n(S)$ is normalized in range $[0, 1]$ as a final news ranking score. The high quality scores observed in Table 8 demonstrates that we agree with the experts that these sources are important, not just their relative rankings as measured by Spearman correlation.

As shown in Table 8, Of the 1051 distinct sources mentioned in these rankings, 914 (87%) are tracked in *MediaRank*. Fully 34/35 experts exhibit a positive correlation with our rankings. The average Spearman coefficient is 0.57, and average ranking quality score is 0.58. For rankings with p-value < 0.05 (24 rankings marked blue), the average Spearman coefficient is 0.69, and average ranking quality score is 0.63.

External Rankings	News Group		T_n/N	Compared to *MediaRank*		
	Topic	Lang/Nation		Corr.	p-value	Quality
NuzzelRank	All	All	97/99	0.55	3.7E-09	0.87
OnlineCollegeCourse	General	English	10/10	0.28	4.3E-01	0.74
Forbes	General	U.S.	12/12	0.56	5.6E-02	0.75
JournaWiki	General	U.S.	41/42	0.68	1.1E-06	0.55
Ranker	General	U.S.	49/49	0.40	4.2E-03	0.55
FeedSpot U.S.	General	U.S.	97/104	0.95	6.1E-49	0.66
AllYouCanRead U.S.	General	U.S.	28/30	0.60	7.9E-04	0.87
FeedSpot Italian	General	Italian	5/9	0.50	3.9E-01	0.06
AllYouCanRead Italian	General	Italian	29/30	0.39	3.4E-02	0.40
Agility PR Solution	General	Canadian	10/10	0.41	2.4E-01	0.41
FeedSpot Canadian	General	Canadian	57/64	0.79	1.6E-13	0.72
AllYouCanRead Canadian	General	Canadian	30/30	0.72	8.3E-06	0.77
BlogHub	General	French	20/20	0.29	2.1E-01	0.72
FeedSpot French	General	French	8/9	0.55	1.6E-01	0.11
AllYouCanRead French	General	French	29/30	0.64	1.6E-04	0.73
DeutschLand	General	German	4/6	1.00	0.0E+00	0.32
FeedSpot German	General	German	27/30	0.75	6.2E-06	0.61
AllYouCanRead German	General	German	12/14	0.21	5.1E-01	0.57
FeedSpot Spanish	General	Spanish	5/17	0.70	1.9E-01	0.44
AllYouCanRead Spanish	General	Spanish	19/30	0.78	9.6E-05	0.53
FluentU	General	Russian	3/7	-0.50	6.7E-01	0.70
FeedSpot Russian	General	Russian	6/9	0.94	4.8E-03	0.18
AllYouCanRead Russian	General	Russian	27/30	0.52	5.5E-03	0.48
Penceo Sport	Sport	All	12/15	0.84	6.4E-04	0.64
FeedSpot Sport	Sport	All	30/52	0.65	8.8E-05	0.45
AllYouCanRead Sport	Sport	All	20/24	0.66	1.6E-03	0.80
MakeUseOf	Entertain	All	8/10	0.36	3.9E-01	0.47
FeedSpot Entertain	Entertain	All	13/22	0.18	5.7E-01	0.16
AllYouCanRead Entertain	Entertain	All	20/24	0.51	2.1E-02	0.88
eBizMBA	Business	All	11/15	0.75	8.5E-03	0.78
FeedSpot Business	Business	All	39/46	0.88	8.3E-14	0.52
AllYouCanRead Business	Business	All	25/26	0.49	1.4E-02	0.87
WebTopTen	Tech	All	10/10	0.68	2.9E-02	0.76
FeedSpot Tech	Tech	All	69/84	0.91	1.4E-26	0.55
AllYouCanRead Tech	Tech	All	32/32	0.37	3.6E-02	0.74

Table 8: Comparisons of *MediaRank* to 35 expert news rankings. "Quality" measures the normalized *MediaRank* scores of common sources, with range [0, 1]. 24 rankings are above 0.05-significance level. Their average Spearman coefficient is 0.69, and average ranking quality score is 0.63.

Table 9 shows presents the top ten news sources by *MediaRank* in each of five topic domains. The sources that also appear on *NuzzelRank*'s top 99 list are highlighted in bold. There is general agreement between the two systems, particularly among General, Business and Technology. The range of signal ranking percentile is [0.0001, 1]. The smaller the percentile value is, the better quality a source has regarding the signal. Bias is assigned zero for non-political news. We can see that the lower ranking news have darker color. The *Daily Mail* has large breadth, reputation and popularity scores, but its ranking is downgraded due to aggressive ads display.

9 CONCLUSIONS

We have demonstrated that the quality of news sources can be instructively measured using a mix of computational signals reflecting the peer reputation, reporting bias, bottomline pressure, and popularity. Our immediate focus now revolves around engineering improvements to our article analysis, such as improved non-English language support for political bias measurement, e.g.

General	Sport	Business	Entertainment	Technology
nytimes.com	**espn.com**	**bloomberg.com**	hollywoodreporter.com	**theverge.com**
washingtonpost.com	sbnation.com	**wsj.com**	**variety.com**	**techcrunch.com**
theguardian.com	mlb.com	**businessinsider.com**	people.com	**wired.com**
cnn.com	nfl.com	**forbes.com**	deadline.com	**recode.net**
bbc.com	fifa.com	**cnbc.com**	tmz.com	**cnet.com**
reuters.com	si.com	**ft.com**	ew.com	**arstechnica.com**
usatoday.com	skysports.com	**hbr.org**	billboard.com	**engadget.com**
politico.com	baseball-reference.com	**marketwatch.com**	**rollingstone.com**	**zdnet.com**
npr.org	cbssports.com	fool.com	**vanityfair.com**	autonews.com
dailymail.co.uk	rotoworld.com	investopedia.com	pagesix.com	autocar.co.uk

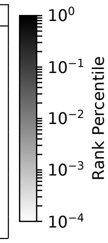

Table 9: *MediaRank* top 10 news of different topics. Sources ranked top in *NuzzelRank* are shown in bold. Strong agreement in "General", "Business" and "Technology". The rank percentiles of six signals are also visualized (from left to right: reputation, popularity, breadth, bias, social bot and ads scores). Lower ranking sources have lower ranking signals, thus marked in darker color. The *Daily Mail* has strong breadth signal, but it is downgraded due to aggressive ads display.

Russian, Chinese and Japanese. We are also working on improved visualization techniques for news analysis, to be reflected at www.media-rank.com.

Deeper NLP analysis of articles to verify or dispute factual claims is a longer-term goal of this work. The data collected and released over the course of our *MediaRank* project will be a valuable asset to such work.

ACKNOWLEDGMENTS

We thank Prof. Michael Ferdman's help for setting up the server cluster. We are grateful for efforts of Charuta Pethe, Ankur Rastogi, Mohit Goel, Harsh Agarwal, Rohit Patil, Abhishek Reddy in building and maintaining the demonstration website. This work was partially supported by NSF grant IIS-1546113. Any conclusions expressed in this material are of the authors' and do not necessarily reflect the views, either expressed or implied, of the funding party.

REFERENCES

[1] Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angelo Spognardi, and Maurizio Tesconi. 2015. Fame for sale: efficient detection of fake Twitter followers. *Decision Support Systems* 80 (2015), 56–71.

[2] Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angelo Spognardi, and Maurizio Tesconi. 2017. The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race. In *Proceedings of the 26th International Conference on World Wide Web Companion*. 963–972.

[3] Gianna M Del Corso, Antonio Gulli, and Francesco Romani. 2005. Ranking a stream of news. In *Proceedings of the 14th international conference on World Wide Web*. ACM, 97–106.

[4] Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd annual meeting on association for computational linguistics*. 363–370.

[5] Seth Flaxman, Sharad Goel, and Justin M Rao. 2016. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly* 80, S1 (2016), 298–320.

[6] CJ Hutto Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International Conference on Weblogs and Social Media*.

[7] Namrata Godbole, Manja Srinivasaiah, and Steven Skiena. 2007. Large-Scale Sentiment Analysis for News and Blogs. *International Conference on Weblogs and Social Media* 7, 21 (2007), 219–222.

[8] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 855–864.

[9] Yang Hu, Mingjing Li, Zhiwei Li, and Wei-ying Ma. 2006. Discovering authoritative news sources and top news stories. In *Asia Information Retrieval Symposium*. Springer, 230–243.

[10] Keisuke Kiritoshi and Qiang Ma. 2014. Named entity oriented related news ranking. In *International Conference on Database and Expert Systems Applications*. Springer, 82–96.

[11] Liang Kong, Shan Jiang, Rui Yan, Shize Xu, and Yan Zhang. 2012. Ranking news events by influence decay and information fusion for media and users. In *Proceedings of the 21st ACM international conference on Information and knowledge management*. ACM, 1849–1853.

[12] Vivek Kulkarni, Junting Ye, Steve Skiena, and William Yang Wang. 2018. Multi-view Models for Political Ideology Detection of News Articles. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 3518–3527.

[13] Kyumin Lee, Brian David Eoff, and James Caverlee. 2011. Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter.. In *International AAAI Conference on Weblogs and Social Media*. 185–192.

[14] Xi Mao and Wei Chen. 2010. A method for ranking news sources, topics and articles. In *2nd International Conference on Computer Engineering and Technology*, Vol. 4. IEEE, V4–170.

[15] Amy Mitchell, Katie Simmons, Katerina Masta, and Laura Silver. 2018. Publics Globally Want Unbiased News Coverage, but Are Divided on Whether Their News Media Deliver. *Pew Research Center* (2018).

[16] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. *The PageRank citation ranking: Bringing order to the web.* Technical Report. Stanford InfoLab.

[17] Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. 2018. Automatic Detection of Fake News. In *Proceedings of the 27th International Conference on Computational Linguistics*. Association for Computational Linguistics, 3391–3401.

[18] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 701–710.

[19] Fritz Plasser. 2005. From hard to soft news standards? How political journalists in different media systems evaluate the shifting quality of news. *Harvard International Journal of Press/Politics* 10, 2 (2005), 47–68.

[20] Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein. 2018. A stylometric inquiry into hyperpartisan and fake news. In *ACL*. 231–240.

[21] Natali Ruchansky, Sungyong Seo, and Yan Liu. 2017. Csi: A hybrid deep model for fake news detection. In *Proceedings of CIKM*. ACM, 797–806.

[22] Ana Lucía Schmidt, Fabiana Zollo, Michela Del Vicario, Alessandro Bessi, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. 2017. Anatomy of news consumption on Facebook. *Proceedings of the National Academy of Sciences* 114, 12 (2017), 3035–3039.

[23] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. 2018. The spread of low-credibility content by social bots. *Nature communications* 9, 1 (2018), 4787.

[24] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter* 19, 1 (2017), 22–36.

[25] Alexandru Tatar, Panayotis Antoniadis, Marcelo Dias De Amorim, and Serge Fdida. 2014. From popularity prediction to ranking online news. *Social Network Analysis and Mining* 4, 1 (2014), 174.

[26] Maartje ter Hoeve, Anne Schuth, Daan Odijk, and Maarten de Rijke. 2018. Faithfully Explaining Rankings in a News Recommender System. *arXiv preprint arXiv:1805.05447* (2018).

[27] Laurens Van Der Maaten. 2014. Accelerating t-SNE using tree-based algorithms. *The Journal of Machine Learning Research* 15, 1 (2014), 3221–3245.

[28] Onur Varol, Emilio Ferrara, Clayton A Davis, Filippo Menczer, and Alessandro Flammini. 2017. Online human-bot interactions: Detection, estimation, and characterization. *arXiv preprint arXiv:1703.03107* (2017).

[29] Onur Varol, Emilio Ferrara, Clayton A Davis, Filippo Menczer, and Alessandro Flammini. 2017. Online human-bot interactions: Detection, estimation, and characterization. In *Eleventh International AAAI Conference on Web and Social Media*.

[30] Amy X Zhang, Aditya Ranganathan, Sarah Emlen Metz, Scott Appling, Sehat, et al. 2018. A structured response to misinformation: defining and annotating credibility indicators in news articles. In *Companion of the The Web Conference 2018*. 603–612.