Analysing Biases in Perception of Truth in News Stories and their Implications for Fact Checking

Mahmoudreza Babaei
Max Planck Institute for Software Systems (MPI-SWS)
Germany
babaei@mpi-sws.org

Abhijnan Chakraborty
Max Planck Institute for Software Systems (MPI-SWS)
Germany
achakrab@mpi-sws.org

Juhi Kulshrestha
Max Planck Institute for Software Systems (MPI-SWS)
Germany
juhi@mpi-sws.org

Elissa M. Redmiles
University of Maryland
United States
eredmiles@cs.umd.edu

Meeyoung Cha
Graduate School of Culture Technology, KAIST
South Korea
meeyoungcha@kaist.edu

Krishna P. Gummadi
Max Planck Institute for Software Systems (MPI-SWS)
Germany
gummadi@mpi-sws.org

ABSTRACT

A flurry of recent research has focused on understanding and mitigating the threat of "fake news" stories spreading virally on social media sites like Facebook and Twitter. In this work, we focus on how users *perceive truth* in viral news stories. To this end, we conduct online user-surveys asking people to *rapidly* assess the likelihood of news stories being true or false. Our goal is to quantify the extent to which users can *implicitly* recognize (perceive) the accurate truth-level of a news story (obtained from fact checking sites like Snopes).

Our analysis of users' implicit perception biases (*i.e.,* inaccuracies in estimating truth-level of stories) reveals many interesting trends. For instance, we observe that in the set of stories fact checked by Snopes, the perception biases are not correlated with the actual truth-level of the news stories. Our finding implies that there exist as many true stories that are believed by users to be more false than they actually are, as there exist false stories that are believed to be more true than they actually are. We argue that the stories that are in need of being fact checked are the stories where users exhibit the largest perception biases. However, we show that existing fact checking strategies that rely on users to report stories they suspect to be false, would prioritize fact checking stories based on their actual truth-level rather than perception biases. We propose an alternative strategy to select stories with large perceived biases for fact checking.

CCS CONCEPTS

• **Information systems → Data mining**;

KEYWORDS

Perception bias, False news, Fact-check

FairUMAP2018, July 2018, Singapore

1 INTRODUCTION

Recently social media sites like Facebook and Twitter have been severely criticized by technologists, policy makers, and media watchdog groups for allowing *fake news* stories to spread unchecked on their platforms [4, 8, 9]. In response, Facebook [1] and Twitter [2] have started to encourage their users to report any news story that they encounter on the site, which they perceive as fake. Stories that are reported as fake by a large number of users are prioritized for *fact checking* by (human) experts at fact checking organizations like Snopes [3] and PolitiFact [4], according to the Poynter's Code of Principles [5]. Stories deemed as false by fact checkers are prominently labeled as *disputed* by Facebook. Thus, social media sites' strategies to select a small set of stories to fact check relies crucially on the ability of user crowds to perceive (un)truth in news stories and report them.

However, to date, few studies, have focused on understanding how users perceive truth in news stories or how biases in their perceptions might affect current strategies to detect and label fake news stories. To illustrate the need for fact checking systems to account for users' truth perception biases, in Figure 1, we show how users' perceived truth levels of news stories compare with their ground truth levels. The stories that are likely to be reported (flagged) by most users for fact checking are the following two stories (S_1 and S_2) that have the lowest perceived truth levels:

(1) *False Story* S_1*:* President Trump inherited a White House infested with cockroaches due to the careless behavior of his predecessor, Barack Obama. [6]

(2) *Mostly False Story* S_2*:* President Donald Trump changed the constitution to read 'citizens' instead of 'persons.' [7]

[1] https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/

[2] https://www.washingtonpost.com/news/the-switch/wp/2017/06/29/twitter-is-looking-for-ways-to-let-users-flag-fake-news/

[3] https://www.snopes.com/ratings/

[4] http://www.politifact.com/truth-o-meter/article/2013/nov/01/principles-politifact-punditfact-and-truth-o-meter/

[5] https://www.poynter.org/international-fact-checking-network-fact-checkers-code-principles

[6] https://www.snopes.com/obama-left-trump-a-white-house-full-of-roaches/

[7] https://www.snopes.com/did-trump-alter-the-constitution/

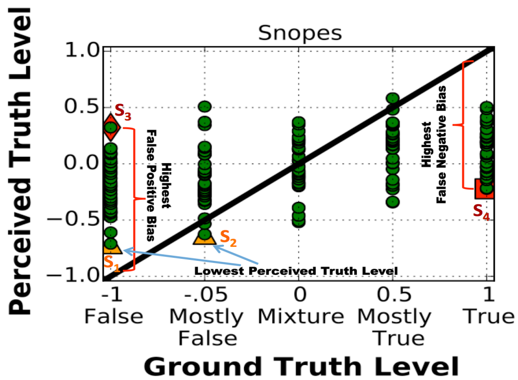


Figure 1: A scatter plot of ground truth levels and perceived truth levels for 150 news stories. Ground truth levels of stories were determined by the Snopes fact checking site, while their corresponding perceived truth levels were estimated by averaging truth level ratings provided by 100 surveyed users. The stories that lie along the diagonal are ones for which users correctly perceived their ground truth value. The stories above the diagonal are perceived to be more true (less false) than they actually are (*i.e.,* false positive bias in perception), while those below the diagonal are perceived to be less true (more false) than they actually are (*i.e.,* false negative bias in perception).

However, note that these stories are already perceived accurately by users to be mostly or completely false. We argue that there is little to be gained by fact checking stories whose truth value is correctly perceived by most users.

On the contrary, consider the following two stories (S_3 and S_4) with the highest false positive and false negative biases in their truth perceptions, respectively:

(3) *False Story* S_3*:* Sen. John McCain's vote against a 'skinny repeal' health care proposal stopped attempts to repeal the Affordable Care Act for FY '17. [8]

(4) *True Story* S_4: The national debt saw a 'surprising' decline of \$102 billion between 20 January and 27 July 2017. [9]

Note that stories S_3 and S_4 are considerably *more damaging* than S_1 and S_2 because the truth values of the former are incorrectly perceived while truth values of the latter are correctly perceived by the readers. Nevertheless, S_1 and S_2 are more likely to be reported by users and fact checked with greater priority than S_3 and S_4! In fact, in todays' social media sites, the higher the false positive bias in the perception of a story, the less likely it is to be reported for fact checking. Worse, a true story like S_4 that is mistakenly perceived by many users as false would not be labeled by current Facebook policies as "disputed", even if it were fact checked, because only false (and not true) stories can be disputed.

Our discussion above highlights the pitfalls of ignoring user truth perceptions when fact checking stories. Specifically, we observe that (i) fact checking has greater utility when it targets stories with high truth perception bias rather than stories that are just false (*i.e.,* low ground truth value) and (ii) relying on users to report stories they perceive as false for fact checking prioritizes easily identifiable fake stories (S_1 and S_2) over harder to identify fake stories (S_3) and hard to believe true stories (S_4).

Against this background, in this paper, we present an in-depth analysis of how users perceive truth in news stories. Specifically, we present (i) an exploratory analysis (characterization) of users' truth perception biases for stories fact checked by Snopes and (ii) a predictive analysis (implications) of the perception biases for leveraging crowd wisdom for fact checking. In the process, we make three primary types of contributions:

1. *Methodological:* We developed a new method for assessing *implicit* truth perceptions of users. Our test, inspired by Implicit Associativity Tests, asks users to *rapidly* assess how truthful or untruthful the claims in a news story are.

2. *Empirical:* Our exploratory analysis of users' truth perception biases yielded several interesting findings. For instance, (i) for many stories, the collective wisdom of the crowds (average truth level ratings) differs significantly from their ground truth level, *i.e.,* wisdom of crowds is inaccurate, (ii) across different stories, the false positive perception bias (*i.e.,* a gullible user perceiving the story to be more true than it is in reality) is as big a concern as the false negative perception bias (*i.e.,* a cynical user perceiving a story to be more false than it is in reality), and (iii) when the truth levels of stories are highly disputed (*i.e.,* show high variance), it is frequently the result of users' political ideologies (*i.e.,* whether they support democrats vs. republicans) influencing their truth perceptions.

3. *Practical:* Our predictive analysis of users' perception biases reveals the limitations of current strategies for selecting a small set of news stories to fact check based on how many users report the story as fake. We argue that an alternate strategy that relies on measuring *disputability* (variance) in truth perceptions of crowds would prioritize stories that suffer from the largest perception biases.

2 METHODOLOGY AND DATASETS

2.1 Designing Truth Perception Tests

The goal of our tests is to gather data about how users implicitly perceive truth in news stories. Our test design is inspired by *Implicit Association Tests* [5] where people are asked to rapidly associate words with different categories to evaluate the strength of association between concepts and people's implicit evaluations.[10]

We performed our *Truth Perception Tests* as Amazon Mechanical Turk (AMT) [1] survey experiments where we showed AMT workers news stories as claims, and asked them to label the claims as either "I can confirm it to be false", "very likely to be false", "possibly false", "can't tell", "possibly true", "very likely to be true", or "I can confirm it to be true".

2.2 Ground Truth Labeled News Stories Datasets

We performed the Truth Perception Tests over ground truth labeled news stories from Snopes, which categorizes news stories into five ground truth labels – False, Mostly False, Mixture, Mostly True, and True.[11] We mapped these ground truth labels on a scale between

[8] https://www.snopes.com/mccains-vote-obamacare-repeal/
[9] https://www.snopes.com/national-debt-trump/
[10] https://implicit.harvard.edu/implicit/iatdetails.html
[11] https://www.snopes.com/

-1.0 (False) and +1.0 (True). We collected the 30 most recently fact checked news stories (which have been categorized by Snopes as Politics) from each ground truth label, getting a total of 150 news stories in our dataset.

2.3 Gathering Users' Implicit Truth Perceptions

By conducting the Truth Perception Tests as AMT surveys, we gathered the truth perceptions of 100 AMT master workers [1] from the US for each news story in our dataset. We observed that on an average the AMT workers took 11 seconds to rate their perceived truth levels of a story. This observation confirms that users gave rapid responses to our tests to measuring their implicit truth perceptions, which is a hallmark of implicit tests.

We also asked the workers for their demographic information including their political leanings. Out of the workers who took our tests, 53.7% were democrats, 20.4% republicans, 21.6% neutral and 4.2% did not disclose their leaning. While the demographic distributions of workers may not be representative of the offline population, we can still draw many important observations from this data. In the future, we plan to repeat the experiments with demographically representative set of users using the participant pool of a US survey company to overcome these limitations and to also study the impact of users' other demographic characteristics on their perceptions of truth in news stories.

3 MEASURES OF PERCEPTION BIAS

By perception bias (PB) of a user U for a news story S, we refer to the error or deviation between the ground truth level (GTL) of the story S and the user U's perceived truth level (PTL_U) of the story S. Therefore, for each story we have two associated truth levels:

- *Ground Truth-Level (GTL):* It is given by the ground truth labels for news stories in the dataset and takes a value between -1.0 and + 1.0. The closer the GTL is to -1.0, the more false the story has been labeled and the closer it to +1.0, the more true the story has been labeled by the fact checking websites.
- *Perceived Truth-Level (PTL):* It is the aggregated value of individual users' truth perceptions (PTL_u) for a story S, and is given by:

$$\text{PTL(S)} = \frac{\sum_{u=1}^{N} PTL_u(S)}{N} \tag{1}$$

where N is the total number of users whose truth perceptions for the story S are being aggregated. The closer the value of PTL(S) is to -1.0, the more the users perceive story S to be false and the closer it is to +1.0 the more the users perceive it to be true.

Based on these truth levels of each story, we compute the following measures to aggregate the individual perception biases of a set of users for each news story:

- **Mean Perception Bias (MPB)** of a story measures the error in the collective perceptions of users (*i.e.,* wisdom of the crowds) in assessing the truth level of a story. Therefore, the Mean Perception Bias for a story S is given by:

$$\text{MPB(S)} = PTL(S) - GTL(S) \tag{2}$$

- **False Positive Bias (FPB)** of a story S measures the *gullibility* of users in their perception of the truth level of the story, *i.e.,* how much the users have over-estimated the truth level of the story by rating it to be more true than it is according to ground truth. False Positive Bias of a story S is computed as follows:

$$\text{FPB(S)} = \begin{cases} \dfrac{\sum_{u=1}^{N_{gullible}} (PTL_u(S) - GTL(S))}{N_{gullible}}, & \text{when } PTL_u\text{(S)} > \text{GTL(S)} \\ 0, \text{otherwise} \end{cases} \tag{3}$$

Here $N_{gullible}$ is the number of gullible users, *i.e.,* users whose perceived truth level ($PTL_u(S)$) is greater (more true) than the ground truth level (GTL(S)) of the story.

- **False Negative Bias (FNB)** of a story S measures the *cynicality* of users in their perception of the truth level of the story, *i.e.,* how much the users have under-estimated the truth level of the story by rating it to be less true than it is according to the ground truth. False Negative Bias of a story S is computed as follows:

$$\text{FNB(S)} = \begin{cases} \dfrac{\sum_{u=1}^{N_{cynical}} (GTL(S) - PTL_u(S))}{N_{cynical}}, & \text{when } PTL_u\text{(S)} < \text{GTL(S)} \\ 0, \text{otherwise} \end{cases} \tag{4}$$

Here $N_{cynical}$ is the number of cynical users, *i.e.,* users whose perceived truth level ($PTL_u(S)$) is lesser (more false) than the ground truth level (GTL(S)) of the story.

- **Total Perception Bias (TPB)** of a story S captures the total error (gullibility or cynicality) in the users' perceptions of truth levels of the story, and is given by

$$\text{TPB(S)} = \frac{\sum_{u=1}^{N} |PTL_u(S) - GTL(S)|}{N} \tag{5}$$

where N is the total number of users whose truth perceptions of the story S are being aggregated.
It is worth noting that while the errors in truth perceptions of gullible and cynical users can counteract one another in the case of MPB, the two errors in perception add up in determining TPB.

- **Variance in Perception Biases (VPB)** of a story S captures the *disputability* in users' truth perceptions of the story and is measured as follows:

$$\text{VPB(S)} = \frac{\sum_{u=1}^{N} (PTL_u(S) - PTL(S))^2}{N} \tag{6}$$

where N is the total number of users whose truth perceptions of the story S are being aggregated. The higher the value of VPB, the more the disagreement or dispute amongst the users about the story's truth level.

4 ANALYZING TRUTH PERCEPTION BIASES

In this section, we use the previously defined measures of perception bias to investigate the truth perceptions of news stories along three dimensions: (i) bias in *collective* truth perceptions, *i.e.,* aggregated *wisdom of crowds*, (ii) bias in *individual* truth perceptions, and (iii) *disputability* of individual truth perceptions.

4.1 Bias in Collective Wisdom of Crowds

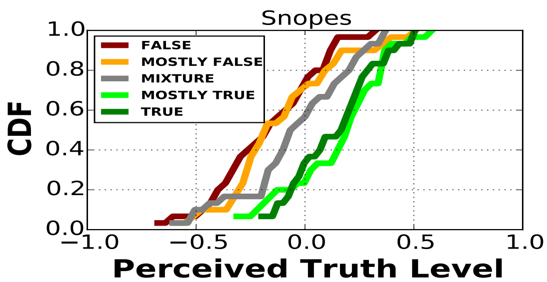


Figure 2: CDF of the Perceived Truth Levels (PTL) for the Snopes news stories with different ground truth levels.

Our investigation of the biases in collective truth perceptions (wisdom of crowds) is motivated by two high-level questions:

(1) Can the wisdom of crowds be leveraged for assessing the truth level of news stories (*i.e.,* fact checking stories)?
(2) Is wisdom of crowds better at assessing the truth levels of true stories or false stories?

	Truth-Levels	PTL True	PTL False
Snopes	GTL True	**75%**	25%
	GTL False	32%	**68%**

Table 1: Comparison of ground truth-levels (GTL) and perceived truth-levels (PTL) of the news stories in the dataset. Here we consider a positive PTL (GTL) value to indicate the perception (ground truth) to be true and negative PTL (GTL) value to indicate the perception (ground truth) to be false.

To check whether collective truth perceptions are useful for assessing the truth levels of news stories, we begin by plotting the cumulative distribution of perceived truth levels (PTL) for Snopes stories with different ground truth levels (shown in Figure 2). The figure shows a high range in perceived truth levels for all stories, independent of their ground truth level – *e.g.,* PTL values for true stories range from -0.22 to +0.5, while PTL values for false stories range from -0.71 to +0.32. While the distributions of perceived truth levels are different for false and true stories, they also exhibit a significant overlap – *e.g.,* 34% of 'true' and 24% of 'mostly true' stories have negative PTL values, while 27% of 'false' and 30% of 'mostly false' stories have positive PTL values.

Our data suggest that while the collective wisdom of crowds has some predictive power in estimating the ground truth labels, their accuracy would be limited. Table 1 shows the limitations of using PTL values to predict GTL values for news stories at a coarse granularity. The table shows that the percentage of stories for which GTL and PTL values have same signs is 75% or lesser.

4.2 Bias in Individual Perceptions

We next shift our focus to the bias in the individual truth perceptions, which can be of two types: (i) False Positive Bias where gullible users perceive the news stories to be more true (*i.e.,* positive) than their ground truth levels and (ii) False Negative Bias where cynical users perceive the news stories to be more false (*i.e.,* negative) than their ground truth levels. The cynicality (FNB) and gullibility (FPB) combine together to give the Total Perception Bias (TPB) values for the news stories which capture the total harm that the users suffer in terms of how far are their individual truth perceptions from the ground truth levels.

We begin by examining the relative contributions of FNB and FPB to the TPB of a story, and we end by comparing how the stories with different ground truth levels are impacted by these biases.

When we look at the distribution of the Total Perception Bias of all the stories in Figure 3 (red curve), we observe that a substantial fraction of stories has a perception bias of more than 0.5. To examine what is leading to this high TPB, we separately considered the cases of FPB and FNB, to determine whether one dominates the other. Figures 3 (orange and cyan curves) shows that there are many stories which have FPB and FNB values higher than 0.5, indicating that FPB is as large a concern as FNB.

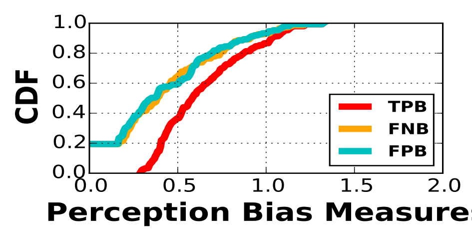


Figure 3: CDF of the Total Perception Bias (red curve), False Negative Bias (orange curve), and False Positive Bias (cyan curve) of news stories.

Examining our stories in more detail, we find that even though false positive and false negative biases may affect very different sets of stories, their overall impact on the total perception bias is comparable across all stories. We argue that high TPB (*i.e.,* high FNB and high FPB) stories should be prioritized for fact checking, since users make the largest errors in judging the truth values of such stories. However, we will show in the next section that current mechanisms for selecting stories for fact checking risk ignoring many of these high TPB stories.

4.3 Disputability in Individual Perceptions

In this section, we analyze the disputability of news stories, *i.e.,* variance in the individual truth perceptions of users. Our analysis is motivated by two high-level questions: (i) How disputed are the truth perceptions of news stories and are true stories more or less disputed than false stories? And (ii) Are highly disputed stories the result of truth perceptions influenced by users' political ideologies?

4.3.1 Significant variance in perception biases for many stories. When we compare the distributions of disputability for stories with different ground truth levels (shown in Figure 4), we observe that their disputabilities have similar distributions, *i.e.,* true stories are

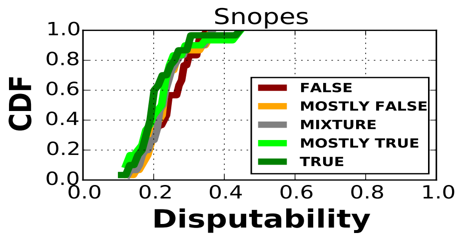
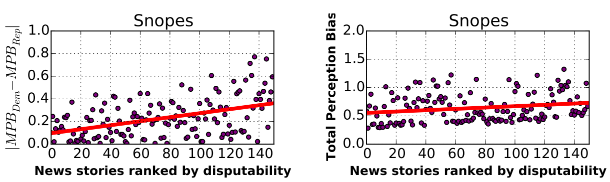


Figure 4: Distribution of disputability (variance in perception bias) of news stories across different ground truth levels.

as disputed as false stories. Thus, we observe that not only are the individual user perceptions of a story biased (as shown in earlier sections), but the biases also vary significantly between users.

4.3.2 High disputability arises from ideologically polarized perceptions. We now examine if high disputability arises from users of different political ideologies (*e.g.,* democrats and republicans) perceiving truth in news stories differently, *i.e.,* they trust or distrust stories that confirm or contradict their political beliefs. To capture such perception biases, we define a new measure called *Ideological Perception Bias* that is computed as the absolute difference between the Mean Perception Biases of Democrat-leaning and Republican-leaning users (*i.e.,* $|MPB_{Dem} - MPB_{Rep}|$).



(a) (b)

Figure 5: Comparing (a) Ideological Perception Bias ($|MPB_{Dem} - MPB_{Rep}|$) and (b) Total Perception Bias with disputability of news stories from our dataset.

To understand the extent to which truth perceptions in highly disputed stories are ideologically polarized, in Figure 5(a), we plot the ideological perception bias ($|MPB_{Dem} - MPB_{Rep}|$) for news stories in our dataset ranked in the increasing order of their disputability. We observe that disputability is strongly correlated with ideological perception biases. Many stories with high disputability also exhibit high ideological perception biases between Democrats and Republicans and vice-versa.

Our findings here make the case for fact checking stories with high disputability – they are the stories where users' truth perceptions reveal a high degree of ideological polarization. Fact checking and labeling such stories could help establish ground truth for stories whose truth values are most contested.

5 IMPLICATIONS FOR FACT CHECKING

In this section, we explore the implications of our findings about users' perception biases on the fact checking of news stories. Today, fact checking sites like Snopes and PolitiFact employ human experts to assess the truth-level of news stories. While the experts can produce high quality and reliable trust assessments, [12] they are limited in terms of the number of stories they can fact check. To select a small set of stories for fact checking, social media sites like Facebook and Twitter rely on their users to report stories that they suspect to be fake. Our goal in this section is to analyze how the *wisdom of the crowds, i.e.,* users' truth perceptions, can be effectively leveraged to select a small set of stories for experts to fact check.

5.1 Criteria and Strategies for Selecting Stories

We begin with a discussion of the desirable criterion for selecting a small set of stories to fact check.

Today, social media sites like Facebook and Twitter do not *explicitly* define a clear objective criterion for their fact checking. Instead, their criterion is *implicitly* defined by their deployed story selection strategy. Given that their strategies rely on users to report fake news stories, only users who perceive the stories to be false would report them, while users who perceive the stories to be true would remain silent. If we assume that the probability of a user reporting a story would be proportional to the magnitude of her negative perceived truth levels (*i.e.,* the more negatively a user perceives a story, the higher the probability of the user reporting it) [13], then it can be observed that the implicit criterion for currently deployed story selection strategies is their **low perceived truth level**.

With our analysis in the previous section, we argued for two alternative criteria for selecting stories to fact check: (a) **high total perception bias** as it captures the total harm or error in a set of users' truth perceptions of a story, independently of the directionality of the error , and (b) **high ideological perception bias** as it captures how polarized the truth perceptions of a story are in the population.

We now discuss strategies for selecting stories according to the above criteria. To select stories with high ideological perception bias, ideally, we would need background knowledge about the political leanings of the users. However, as we observed in the previous section (Figure 5(a)), disputability (*i.e.,* variance in perception biases) is highly correlated with ideological perception bias and disputability can be computed from users' truth perceptions without any knowledge of their political leanings. So a practical strategy would be to use stories' disputability as a proxy measure for their ideological perception bias while selecting stories for fact checking.

Selecting stories with high total perception bias poses a similar practical problem as it assumes knowledge of ground truth levels of stories (beyond their perceived truth levels). To check if a story's disputability can be used as a proxy measure for their total perception bias, in Figure 5(b), we plotted the Total Perception Bias of news stories in our dataset, where the stories are ranked in the increasing order of their disputability. We notice that stories with low TPB generally tend to have lower disputability and high TPB generally tends to correspond to high disputability. This correlation, while not very strong, suggests disputability of a story might be a viable proxy measure for total perception bias while selecting stories to be fact checked. So, our proposed new strategy for selecting

[12] Fact checkers are bound by *Poynter Code of Principles*: https://www.poynter.org/international-fact-checking-network-fact-checkers-code-principles
[13] Similar assumptions have been made in recent research studies [6]

stories to fact check would rely on ordering stories based on their disputability, *i.e.,* variance in perception biases.

5.2 Comparing Story Selection Strategies

We now compare how the stories selected by current strategy (based on lowest perceived truth level) differ from the stories selected by our new strategy (based on highest disputability).

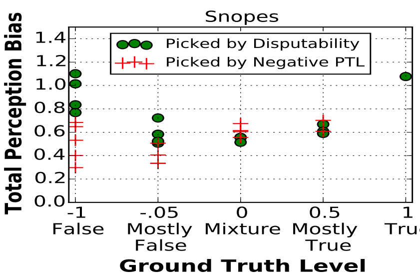


Figure 6: Green dots and red pluses indicate the Total Perception Bias of top 10% news stories ranked by disputability and negative Perceived Truth Values, respectively.

To compare the two strategies, we ranked the stories according to the strategies, and plotted the ground truth levels and total perception biases for the top-10%[14] of the stories in Figures 6. The figures show that (i) current strategies that rely on perceived truth levels tend to pick very few true stories, even when they suffer high total perception bias, while our new disputability-based strategy picks stories from across all ground truth-levels, and (ii) the total perception bias (error in user perceptions) of stories selected by our new strategy tend to be considerably higher than those selected by current strategies.

Note that to estimate disputability of a story; our strategy would require sites to ask a random sample of news readers to provide their truth perceptions. Since the sites only need to ask readers for a rapid truth assessment (*e.g.,* via a prompt on their screen when they read a story that takes only a few seconds to respond), we think this strategy may be viable, in practice. While our strategy is in sharp contrast to current strategies, where sites rely on users to report a story as false, we envisage scenarios where site operators might combine both methods – for example, in a two-stage process, where user reports are used to select a large candidate pool of stories in the first stage, and our disputability measure is used to priority rank stories within the candidate pool in the second stage.

6 RELATED WORK

Over the recent years, a growing amount of effort has been made towards detecting false information (*misinformation* and *disinformation*) by analyzing large-scale digitally logged user behavioral and social network data on the web. There are mainly two lines of research in this direction:

(1) The first line of research has investigated detecting *rumors*, a term used to describe claims that are yet to be verified as 'true' [9, 14]. Based on theoretical studies on characterizing online rumor behaviors [12, 13], computer science researchers have developed multiple rumor detection algorithms using features across multiple categories [8, 10, 11, 16, 18].

(2) The second line of research has specifically focussed on detecting *"fake news"*, which is defined as news that is intentionally and verifiably false [17]. Detecting news articles that contain false claims is a challenging task because human evaluators have shown only marginal improvements (66%) over random guesses (50%) in a recent crowdsourced study [7]. As a preliminary step, recent studies have focused on a fake news problem known as *clickbait articles* or *stance detection*, where news headline and the associated body text have a discordant relationship [3], with methods being developed for detecting such articles using SVM model [2] or neural network based [15] approaches.

The above studies assume fake news stories have been clearly labeled, either by reputable sources or by a large number of individuals, but no study has examined how false or true news stories may be *differently perceived* by people irrespective of their ground truth. Furthermore, these studies do not consider a case when a true story is perceived as false by the audience and hence misses the chance to propagate properly. In this paper we take a step back and ask whether and how veracity of news stories relate to their perceived quality by the public and measure to what extent false stories get perceived as true and vice versa.

7 CONCLUSION

In this paper, we have deeply examined how users perceive truth in news stories by developing a new method for measuring the *implicit* truth perceptions of users where users are asked to rapidly assess how true or false the claims in a news story are. Our analysis of users' truth perception biases revealed that (i) the collective perceptions or wisdom of crowds is not very accurate in assessing the truth levels of a news story, (ii) the false positive perception bias is as big a concern as false negative perception bias, (iii) when the users' perceptions of the truth levels of a story are highly disputed, then it is frequently because the user's political ideologies influence their perceptions of truth. Finally, our predictive analysis exposes the limitations of current strategies for selecting a small set of stories to fact check by relying on how many users report the story to be fake and we propose an alternative approach of relying on measuring the disputability in truth perceptions of users to select the stories which have largest perception biases.

Our findings can help inform the design of selection mechanisms for stories to fact check and help social media platform providers and fact checking organizations to combat fake news in a cheaper manner. Also, our study sets the stage for an interesting line of research we plan to pursue in the future – to study the impact of users' demographic characteristics on their perception of truth in news.

REFERENCES

[1] The Mechanical Turk Blog 2011. Get Better Results with Less Effort with Mechanical Turk Masters, http://mechanicalturk.typepad.com/blog/2011/06/get-betterresults-with-less-effort-with-mechanical-turk-masters-.html.

[14] Put differently, these stories would be prioritized for fact checking over other stories in the dataset according to the two strategies.

[2] Abhijnan Chakraborty, Bhargavi Paranjape, Sourya Kakarla, and Niloy Ganguly. 2016. Stop clickbait: Detecting and preventing clickbaits in online news media. In *Proceedings of the ASONAM*.
[3] Yimin Chen, Niall J Conroy, and Victoria L Rubin. 2015. Misleading online content: Recognizing clickbait as false news. In *Proceedings of the ACM Workshop on Multimodal Deception Detection*.
[4] Adrien Friggeri, Lada A Adamic, Dean Eckles, and Justin Cheng. 2014. Rumor Cascades.. In *ICWSM*.
[5] Anthony G. Greenwald, Debbie E. McGhee, and Jordan L. K. Schwartz. 1998. Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology 74.6: 1464âĂŞ1480* (1998).
[6] Jooyeon Kim, Behzad Tabibian, Alice Oh, Bernhard Schölkopf, and Manuel Gomez-Rodriguez. 2017. Leveraging the crowd to detect and reduce the spread of fake news and misinformation. *arXiv preprint arXiv:1711.09918* (2017).
[7] Srijan Kumar, Robert West, and Jure Leskovec. 2016. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *Proceedings of WWW*.
[8] Sejeong Kwon, Meeyoung Cha, and Kyomin Jung. 2017. Rumor detection over varying time windows. *PloS one* 12, 1 (2017), e0168344.
[9] Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. 2013. Aspects of rumor spreading on a microblog network. In *Proceedings of the SocInfo*.
[10] Sejeong Kwon, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. 2013. Prominent features of rumor propagation in online social media. In *Proceedings of the ICDM*.
[11] Jing Ma, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J Jansen, Kam-Fai Wong, and Meeyoung Cha. 2016. Detecting Rumors from Microblogs with Recurrent Neural Networks. In *Proceedings of the IJCAI*.
[12] Jim Maddock, Kate Starbird, Haneen J Al-Hassani, Daniel E Sandoval, Mania Orand, and Robert M Mason. 2015. Characterizing online rumoring behavior using multi-dimensional signatures. In *Proceedings of the CSCW*.
[13] Onook Oh, Manish Agrawal, and H Raghav Rao. 2013. Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. *Mis Quarterly* 37, 2 (2013).
[14] Vahed Qazvinian, Emily Rosengren, Dragomir R Radev, and Qiaozhu Mei. 2011. Rumor has it: Identifying misinformation in microblogs. In *Proceedings of the EMNLP*.
[15] Md Main Uddin Rony, Naeemul Hassan, and Mohammad Yousuf. 2017. Diving Deep into Clickbaits: Who Use Them to What Extents in Which Topics with What Effects? *arXiv preprint arXiv:1703.09400* (2017).
[16] Natali Ruchansky, Sungyong Seo, and Yan Liu. 2017. CSI: A Hybrid Deep Model for Fake News Detection. In *Proceedings of the CIKM*.
[17] 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.
[18] Zhe Zhao, Paul Resnick, and Qiaozhu Mei. 2015. Enquiring minds: Early detection of rumors in social media from enquiry posts. In *Proceedings of the WWW*.