

Analysis Design Study for Fake News Identification and Evaluation



Lee Won Park and Hangbae Chang

Abstract With the spread of the Internet and increasing amounts of self-proclaimed journalists, articles both true and inaccurate fill the web. These inaccurate articles, most commonly referred to as fake news, have proved to spread quickly and have immense social influence in society. Attempts to detect fake news articles through deep learning techniques and artificial intelligence prove the challenges in fake news detection. While detection techniques are still in development, there is not much research on how readers can discern fake news without technological aid. This paper addresses such limitations regarding the study of fake news detection and provide a detection model for readers. The model is based on logical steps built on detection cues mentioned in previous works. The appropriateness of the detection cues will be determined based on case studies.

Keywords Fake news · Fake news identification · Industrial security

1 Introduction

Fake news has always been around, and always has been a societal issue. However, the US presidential election in 2016 changed the public's perception of fake news. Previously, fake news was frequently referred to news parody or satirical news. The core distinction between fake news and parody or satirical news is that the former has the clear intention to deceive people, whereas the latter intends to make open jokes without the intention to provide false information. Additionally, previous works have proven that those who are familiar with this form of news parody or satire tend to be more politically informed and interested, compared to those who do not follow such commentary [1].

L. W. Park
Department of Security Convergence, Graduate School, Chung-Ang University,
Seoul 06974, Korea

H. Chang (✉)
Department of Industrial Security, Chung-Ang University, Seoul 06974, Korea
e-mail: hbchang@cau.ac.kr

Fake news, however, can be defined as false news that aims to deceive readers and audiences. Fake news is meant to deceive, and thus often difficult to prove false. Although some believe that fake news is easily discernible, it has been proven that some people are especially susceptible to fake news because they strongly relate with topics of fake news. These groups of individuals also risk not being able to objectify the news they consume because they remain ingrained in their own social bubble, without finding the need to explore whether the news they consume is true or false [2]. However, certain social groups are not the only targets of fake news. Previous works have proven that individuals generally detect fake news correctly only half of the time—that is slightly above chance [3].

Previous works have made efforts to detect fake news based on the sentence structure, vocabulary, source of news, structure of news links, etc. Although the use of AI does aid the detection of fake news across the web, thus providing the potential to stop the initial spread of fake information, it does little to aid the detection process for individuals every day. Part of stopping the initial spread of fake news has to do with guiding individuals to also be able to detect fake news, and thus prevent them from sharing and thus participating in the spread of the fake news.

Therefore, this research is dedicated to exploring the methods by which people can easily detect fake news based on notable cues and logical steps. The research will present cues that will help people notice or at least doubt whether the news article is fake or not and observe whether the cues are applicable in real fake news articles.

2 Previous Works

Previous works surrounding fake news have often been centered around defining what fake news may incorporate, what forms fake news may take, and what fake news looks like in different forms of news media.

Previous works note that satire intentionally utilizes cues to reveal its own deceptiveness, since satirical news is not used as a method to deceive with malintent [4]. The research builds on previous work in satire detection to propose an SVM-based algorithm enhanced with 5 predictive features; grammar, humor, negative affect, absurdity, and punctuation. The paper also agrees with previous works in that people are generally unable to effectively recognize deception in news or in general circumstances.

Similarly, the problem arises when people consume fake news while unaware that it is fake news. Previous works argue that the most significant aspect about contemporary fake news is that it deliberately uses “empathic media”, news that targets the emotions of audiences. The paper notes that fake news can be created to build “fellow-feeling” or group emotion behavior within social networks”, and that fake news concerns “economics of emotion” [5].

Due to the fact that people will half the time or more, fail in discerning between fake and real news, previous studies have also made numerous attempts to provide insight on how to detect fake news via behavioral analysis, syntax, URL structures,

tone or wording of the article, etc. Previous works also provide 9 requirements for fake news detection, those being; the availability of both false and true instances, verifiability of “ground truth”, digital textual format accessibility, homogeneity in lengths and in writing matter, manner of news delivery, predefined time frame, pragmatic concerns, and language and culture [6].

Other research suggests identifying reliable news sources, and assessing the pieces used by the journalist in the piece [7]. Additionally, the research suggests that people should be on the lookout for bias and logical fallacies. The paper classifies bias as something that can easily be identified through the words used in a piece. Since bias is subjective, words that express personal opinions or feelings may be indicative of bias in news articles. The research stresses the importance of looking for facts without any sort of non-objective descriptors.

As fake news is also often classified as clickbait, previous research has also made distinctions between content cues and non-text cues that may help discern clickbait [8]. The paper presents content cues as lexical and semantic levels of analysis, and syntactic and pragmatic levels of analysis. Non-text cues presented in the paper are image analysis and user behavior analysis. The paper notes that images can attract attention and because it can draw visual attention of readers, the image is processed before the article is read. Research notes that clickbait frequently uses images to draw in the interests of users through misinformation, and readers tend to receive spontaneously integrated information from the headline with a photo in the associated article.

3 Materials and Methods

3.1 Model

Based on previous works, this paper will present a logical method that will aid people in the process of discerning fake news online.

News consumers are advised to first consider the title of the news article. Previous works suggest that the title of the article may have different layers of fakeness to it. Such attributes of fake article titles may include (a) overrun sentences, (b) sentences with disparate topics, (c) sentences with unnecessary overuse of adjectives, (d) sentences that create strong emotions such as anger, disgust, hate, or sadness, and (e) topics that are completely novel.

After initially clicking on the article based on the title, readers will then be able to clearly note the source of the article. The URL and source may be quick and accurate benchmarks for readers to note whether the article is reliable or not. Once readers have checked the title and source of the article, they will then dissect the body of the article. The body of the article is analyzed similarly to how the title is analyzed. Since there is more text in the body than the title, there may be more frequent and obvious non-textual and textual cues that indicate the level of fakeness

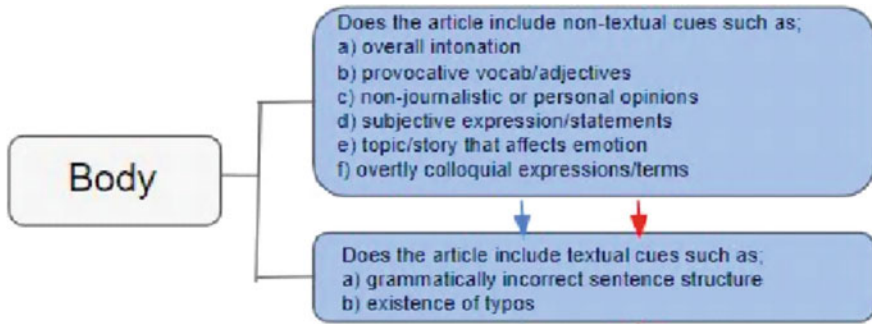


Fig. 1 Discerning fakeness from the body of the article

of the news article. Non-textual cues may include the intonation of the article, the level of humor or satire, the use of inflammatory adjectives, the use of subjective descriptions, the existence of opinion of the writer, etc. Textual cues may include the sentence structure of the article, the grammatical accuracy of the sentences within the article, the existence or frequency of typos, etc.

Finally, readers may be able to discern the level of fakeness of a news article based on the inserted photo in the article. Manipulated photos are of course fake news, but the article can also be suspected as a fake if the inserted photo is completely irrelevant to the title or context of the article (Fig. 1).

Based on the assessment points discussed above, a logical algorithm was developed. However, this paper will only examine and test fake news identification cues via the body of text. The logical algorithm for analyzing the body of text is presented below.

3.2 Case Study Test

Based on the logical detection model suggested above, this paper tests how appropriate and applicable the model is on existing fake news articles. The article presented is primarily judged on the body of the article. The fake news article introduced in the paper is one of the top fake news articles that received the most hits of Facebook in 2018. The article title is “Muslim Figure: “We have Pork-Free Menus Or We Will Leave U.S.” How Would You Respond This?” The text of the article is presented below in italics.

“Muslim migrants have turned into a worldwide issue, and the circumstance deteriorates each day. Our nation has had a considerable measure of Muslim-related issues of late, yet this time settlers went too far. President Donald Trump had the ideal answer for the issue forced by outsiders, however Democrats didn’t bolster him. You won’t think about what occurred next...

Table 1 Identified cues for fake news article

Cues	
Grammatical error	✓
Typo	
Overall undertone	✓
Provocative adjectives/statements	
Non-journalistic/personal statements	
Subjective expressions/descriptions	✓
Novel/emotional topic	
Colloquial terms/statements	

Barack Obama urged Muslims to request their rights. Muslims were informed that America is an astonishing nation loaded with potential outcomes. All things considered, it wasn't some time before settlers set particular prerequisites. Some of these individuals do have issues, and need genuine help. Be that as it may, the vast majority of them are simply crossing the US outskirts to appreciate greatest advantages.

President Trump realized this would transform into a difficult issue, and recommended that our nation bars foreigners from a few Muslim dominant part nations. Tragically, the travel boycott was blocked comfortable start, and Muslim outsiders attacked the nation (Table 1).

America resembles a strong and minding host, however some of its guests are accompanying lethal thoughts in their psyche. They are continually requesting more. This time Muslim settlers made a particular request, and it includes sustenance served in schools.”

The analysis of the fake news article is as follows:

- “Issues of late” is grammatically incorrect [textual cues: grammatical error]
- The term “ideal answer” is subjective [non-textual cues: subjective expression]
- The term “difficult issue” is subjective [non-textual cues: subjective expression]
- The phrase “the travel boycott was blocked comfortable start” is grammatically questionable [textual cues: grammatical error]
- The term “attacked” is both provocative and non-factual [non-textual cues: provocative expression and subjective expression]
- The term “America resembles a strong and minding host” uses subjective descriptive terms [non-textual cues: subjective expression]
- The general undertone of the article is politically provocative, resembles hate speech due to subjective and emotional choice of vocabulary, and attempts to make political statements by introducing particular political figures [non-textual cues: provocative overall undertone of article]

4 Conclusion

The research presented in this paper contributes to the current state of fake news research by providing insight on how readers can initially spot fake news to discern or suspect the integrity of a news article. Technological advancements are key to stopping the rapid spread of fake news, individuals are unable to effectively utilize such technologies every day. This paper recognizes the limitations of previous studies and presents a logical map that will help readers to discern or suspect whether the article they are consuming is fake or true.

Only one fake news article was introduced and analyzed in this paper. The article was selected to demonstrate the typical form of fake news that is difficult to discern at first glance due to its length, subject matter, and seemingly legitimate journalistic format. Future research should be conducted on an extended level with more sample articles provided for analysis.

References

1. Tandoc E Jr, Lim Z, Ling R (2018) Defining “Fake News”. *Digi Journalism* 6(2):137–153
2. Spohr D (2017) Fake news and ideological polarization: filter bubbles and selective exposure on social media. *Bus Inform Rev* 34(3):150–160
3. Frank MG, Feeley TH, Paolantonio N, Servoss TJ (2004) Individual and small group accuracy in judging truthful & deceptive communication. *Group Decis Negot* 13:45–59
4. Rubin, V., Conroy, N., Chen, Y., & Cornwell, S.: Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News. *Proceedings of NAACL-HLT*, 7–17. (2016)
5. Bakir V, McStay A (2018) Fake news and the economy of emotions. *Digit Journalism* 6(2):154–175
6. Rubin V, Conroy N, Chen Y (2015) Towards news verification: deception detection methods for news discourse. <https://doi.org/10.13140/2.1.4822.8166>
7. Ireland S (2018) Fake news alerts: Teaching news literacy skills in a meme world. *Refer Librarian* 59(3):122–128
8. Chen Y, Conroy N, Rubin V (2015) Misleading online content: recognizing Clickbait as “False News”. *ACM Workshop on Deception Detection*