

Governance of Digital Platforms in the Modern Age of AI: How to Preserve the Social Media Platform of the Future from the Widespread of Fake Content

Abdallah Musmar

College of Business, University of South Florida, Tampa, FL 33620, musmar@mail.usf.edu

He Zhang

College of Business, University of South Florida, Tampa, FL 33620, hezhang@usf.edu

Sunil Mithas

College of Business, University of South Florida, Tampa, FL 33620, smithas@usf.edu

Balaji Padamanbhan

College of Business, University of South Florida, Tampa, FL 33620, bp@usf.edu

Motivation, Research Gap, and Research Questions

AI has recently contributed in the generation and distribution of content that would potentially be shared among millions of users around the globe. Among the many benefits that can be realized from this advancement, there is a concern regarding the increasingly difficult task to distinguish real information from fake information. The resulting ambiguity could potentially affect the perception of an end user on how they perceive the real world. On a larger scale, a fake event or a series of fake events can potentially affect the overall perception of reality for a larger community or sub-community of users. One category of fake content that can affect people is the generation of fake news where a news source is incorporating them along with real news that has actually happened in a user's news feed. We distinguish this type of misinformation from other types of misinformation such as rumors, clickbaits, satire, as they do not serve the same purpose. The intention of distributing fake news is to mislead users and create an alternative sense of reality.

The main contributors to the change in news distribution and consumption are social media sites. Many researchers attempted to explain how news travels on social media in a follower-followee network (Bastos et al. 2012). We argue that news articles that are not real can be categorized into news articles that are: fake, rumor, clickbait, misleading information, and satire/comedy. Previous efforts have used a subset of these combinations (Waldrop 2017; Rubin et al. 2015; Tandoc Jr. et al. 2017). There are potentially more categories of noise that can be added to this list.

Although there are many prior research articles that had major contributions in defining fake news, characterizing fake news articles, distinguishing them from other types of disinformation and misinformation, explaining the spread of fake news, and detecting fake news using many different classification approaches, we believe that the effect of fake news on processing real news in a complex system has not been fully investigated in prior literature. By designing the environment to resemble the essence of a digital platform that contains news sharing and user reactions on social media, we aim to find the effect of increasing the number of fake news being induced in the environment on how users react to real news. Therefore, our first research question is investigating this effect and can be formally defined as:

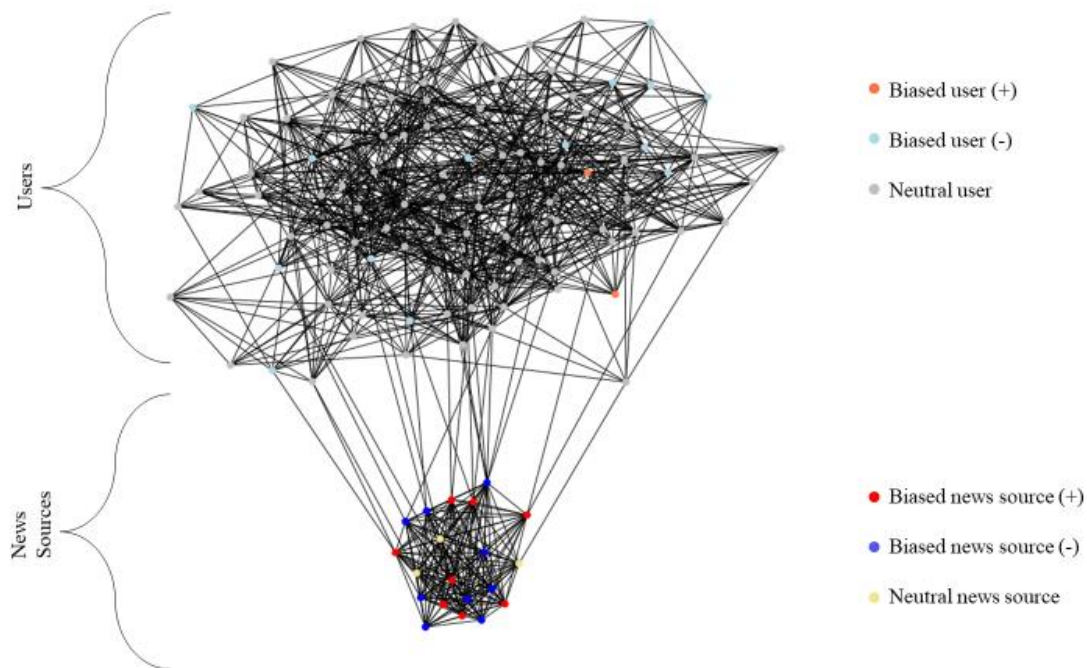
Research Question 1: *What is the effect of injecting “fake news” in a user’s news feed on the users’ decisions when processing “real news”?*

After designing an environment that resembles a digital platform and ensuring that it mimics reality, we want to use the environment as a mechanism to investigate configurations of governance policies that will aid into reducing the effect of fake news on the perception of real news in the network. Therefore, our second research question can be formally defined as:

Research Question 2: *What possible configurations can be implemented on a digital platform in order to reduce the user’s perception of real news as fake news?*

Research Method

The figure below illustrates the concept of recreating a social network in the interest of learning about news events from the users. We have divided the users and news sources each into three types: one neutral type and two types that are biased in opposing directions. In this division we are assuming that all users and news sources are interested in the news topic. Then, we considered the bias of opinion as a main contributor to the structure of the network.



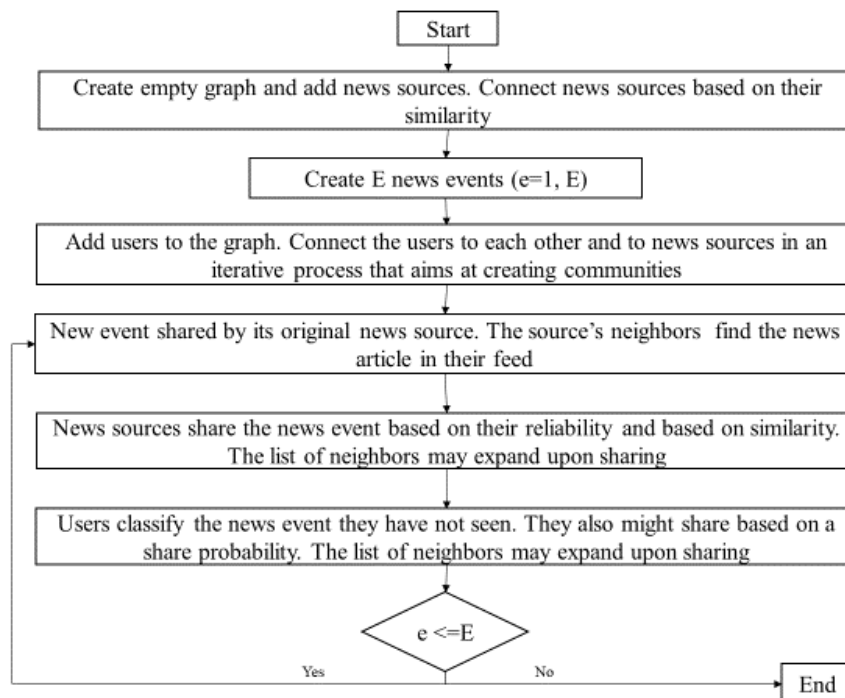
In our simulation design, we can control the percentage at which fake news is being generated. We attempt to gradually increase the number of fake news articles in the environment and observe the number of misclassifications. We expect to find an increasing rate of misclassifications that can be explained by inducing fake news into the environment.

Agent-Based Model Design

The table below contains a summary of the most important features we found in prior literature that would be required to classify a news article as fake or real:

Feature	Feature details
Sentiment of the news article headline	Score is on a scale of (-5,-1) for negative sentiment and (1,5) for positive sentiment
Number of shares	Count of other users who have shared the news article. It must be a natural number.
Source credibility	A percentage that reflects the user's perception about the source of the news article being a reliable reporting agency or not.
Analytical ability and motivation	A binary value where 1 means the user is analytical and motivated, and 0 means the user does not have the ability or motivation.

The flow chart below explains the steps of running our simulation on a high level. We consider each news event generated as a new step in the simulation. :



There are two main decisions the users and news sources make during every step. These decisions are whether they want to spread it or not and their own judgement of the news event that is being shared on their news feed. The decisions made inherently affect other nodes in the network which makes the problem of understanding the mass effect of fake news more complex. However, after a small period of when the step is over, the nature of previously shared news articles is revealed, and it will become common knowledge across the mass. However, we do not expect this knowledge to have an effect on the bias of a user or a news sources at least on the short term. But we do observe the changes in beliefs over many steps and expect a change over a relatively long period of time.

Expected Contributions

In this paper we analyze the relationship of incorporating fake news on the processing of real news. Along with the analysis, we propose a governance framework that will attempt to resolve some of the potential issues that arise from spreading fake information in a digital platform. We think this research contributes to the discussion of discovering the relationship between processing fake news and the accumulated perceived trustworthiness of news. It also provides an in-depth analysis of a complex network structure which is the online news distributors and consumers.

Current Status of the Manuscript

In our paper, we have completed the majority of the introduction, literature review, research questions and hypothesis, and agent-based model design. We are currently working on the governance framework design along with finding meaningful .We are targeting in our submission the Journal of Management Information Systems Special Section on Fake News on the Internet which has a due date of April the 30th, 2020.

References

- Bastos, M. T., Travitzki, R., & Puschmann, C. (2012). *What Sticks With Whom? Twitter Follower-Followee Networks and News Classification*. Association for the Advancement of Artificial Intelligence.
- Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). Deception Detection for News: Three Types of Fakes. *The Annual Meeting of the Association for Information Science and Technology*. S.t. Louis, Missouri.
- Tandoc Jr., E. C., Lim, Z. W., & Ling, R. (2017). Defining “Fake News”. *Digital Journalism*, 6(2), 137-153. doi:10.1080/21670811.2017.1360143
- Waldrop, M. M. (2017). News Feature: The genuine problem of fake news. *Proceedings of the National Academy of Sciences*, 114(48), 12631-12634. doi:10.1073/pnas.1719005114