Modeling Misinformation Diffusion in Social Media: Beyond Network Properties

Francesca Spezzano
Computer Science Department
Boise State University
Boise, ID, USA
francescaspezzano@boisestate.edu

***Abstract*—In this paper, we discuss the current limitations of existing models for misinformation diffusion in social media and present our current work suggesting that other factors beyond network properties play an important node in modeling misinformation spread and profiling fake news spreaders. These factors include news and user characteristics such as user demographics, profile properties, and behavior and activity, and news style and content complexity.**

***Index Terms*—misinformation, fake news diffusion, profiling fake news spreaders, computational social science, diffusion of innovation**

I. Introduction

A democratic and civil society ensures that the public is well informed and can openly share and act on information or opinions [1]. Unfortunately, Web features that allow for openness directly challenge such societal goals as they make it increasingly easy to disseminate misleading stories to manipulate opinions and public response. People have difficulty discerning Web misinformation from truthful information [2], especially teenagers and seniors who are not equipped to spot misleading stories they may read on social media [3], [4]. Moreover, while 64% of Americans believe that misleading stories cause confusion about the basic facts of current events, people frequently share news on social media without even reading what they are forwarding [5]. Misinformation has already affected stock markets [6], slowed response to disasters [7] and terrorist attacks [8], and is threatening public health during the COVID-19 pandemic [9].

Current research is limited in addressing misinformation diffusion and reducing its spread through social media. Existing models for modeling misinformation diffusion assume unknown connections among individuals (implicit network), do not incorporate the cognitive and social processes that people use to consume news [10], [11] and do not incorporate a time dimension modeling when a user will share a fake news item they received on their social media account.

In this paper, we first discuss related work on modeling misinformation diffusion and profiling fake news spreaders and discuss current limitations. Next, we present our ongoing research effort focused on addressing these limitations. Specifically, we propose a user-centric approach based on the diffusion of innovation theory [12], [13] to model the diffusion of fake news in social media and characterize and predict fake news spreaders and their influence. The diffusion of innovation theory considers user and news characteristics important factors to explain news sharing behavior beyond network-related factors. Our research confirms this theory in a real-world Twitter setting.

II. Related Work

It is important to understand the underlying mechanisms modeling its diffusion to prevent misinformation spread in social networks. Related work has proposed stopping the spread of misleading news by blocking its propagation paths, intervening against opinion leaders, or injecting ”corrector users” to counter by spreading true information [14]–[16]. In studying information diffusion, many existing studies have focused on modeling rumor propagation [17], while few have modeled the diffusion of misleading stories. Specifically, these diffusion models for misinformation address the problem similar to the diffusion of an infectious disease (epidemiological models) [18]–[20] or as a Hawkes process [21] and assume unknown connections among individuals (implicit network). This makes these models more suitable to study global patterns, such as trends and ratios of people sharing stories of a given topic, but not the local node to node diffusion patterns. On the other hand, Cho et al. [22] model individual agent's diffusion behavior under information veracity uncertainty and incorporate agent's topic prior belief and neighbors' pairwise interactions. Classical models such as the Independent Cascade and the Linear Threshold models can be used to model misinformation spread in case the user social network is explicit [23]. However, some works went beyond considering just the network to explain news diffusion and tested hypotheses inspired by the Diffusion of Innovation Theory, which also considers user and news characteristics as important factors to explain news sharing behavior [12]. Ma et al. [10] found opinion leadership, news preference, and tie strength to be the most important factors at predicting news sharing, while homophily hampered news sharing in users' local networks. Also, people driven by gratifications of information seeking, socializing, and status-seeking were more likely to share news on social media platforms [24].

Further, to model misinformation spread by non-malicious users who unintentionally engage with it, diffusion models for misinformation spread should incorporate the reader's perspective and also include cognitive and social principles and processes behind how people consume news. Also, the processes that people use to spread misinformation can be different from person to person and also depend on the news domain.

Several recent works have been addressed the problem of profiling fake news spreaders. Vosoughi et al. [25] revealed that fake news spreaders had, on average, significantly fewer followers, followed significantly fewer people, and were significantly less active on Twitter. With respect to the social media platform Twitter, although bots contribute to spreading fake news, the dissemination of fake news on Twitter is mainly caused by human activity.

Shu et al. [26] analyzed user profiles to identify the characteristics of users who are likely to trust or distrust fake news. They found that, on average, users who share fake news tend to be registered for a shorter time than the ones who share real news. Additionally, while bots were shown to be more likely to post a piece of fake news than real news, users who spread fake news are still more likely to be humans than bots. They also show that real news spreaders are more likely to be more popular than fake news spreaders and that older people and females are more likely to spread fake news.

Guess et al. [27] also analyzed user demographics as predictors of fake news sharing on Facebook and found that political orientation, age, and social media usage be the most relevant. In this study, the majority of fake news items included for analysis were dated from 2016 and were typified by pro-Trump sentiments. The researchers found that users who leaned to the political right were more likely to share those fake news items. Additionally, individuals identified as senior citizens tended to share more fake news (a fact the researchers theorized to be due to age-associated lower digital media literacy skills necessary to assess online news truthfulness). Finally, the researchers found that the more news people post on social media, the less likely they are likely to share fake news, an observation theorized to be the case because those users would be more familiar with the platform and what they share.

The author profiling shared task at the PAN 2020 conference focused on determining whether or not the author of a Twitter feed was keen to spread fake news [28]. Participants proposed different linguistic features to address the problem, including (a) n-grams, (b) style, (c) personality and emotions, and (d) embeddings. Among them, our previous work [29] showed that psycho-linguistic and personality features are significantly associated with user sharing behavior.

In order to study the problem of discriminating between fake news spreaders and fact-checkers, Giachanou et al. [30] proposed an approach based on a convolutional neural network to process a user Twitter feed in combination with features representing user personality traits and linguistic patterns used in their tweets.

A. Current Limitations

Based on our review of related literature, existing work on modeling misinformation diffusion in social media suffers from the following limitations:

- *implicit vs. explicit network*: existing models for misinformation spread assume an implicit network among the social network users, hence they cannot be used to predict which users will be influenced by fake news and how a fake news item may propagate in the network.
- *applying social science-based theories to diffusion*: existing social science studies on fake news sharing behavior are not currently incorporated into fake news diffusion models, which makes these models less tailored to model real-world misinformation diffusion.
- *when fake news is shared*: not everyone takes the same time to re-share news. Some users may share it as soon as it is received, others may take time to analyze the received news, while others may never re-share the received news. Hence, considering introducing a user-news-dependent time component in the news diffusion model can make the simulation of news propagation more realistic.

III. Ongoing Work

Our current work focuses on designing solutions that can overcome the limitations mentioned in the previous section. Specifically, we are working on the following two main problems, namely (i) predicting whether an influenced user will share fake news and (ii) characterizing and predicting fake news spreaders and their influence.

A. Modeling Diffusion of Fake News in Social Media

We propose an approach based on the diffusion of innovations theory to model and characterize how fake news is shared in social media [31]. Specifically, we address the following problem: *given that a user u is influenced on some given (real or fake) news item n by at least one of their followees v (i.e., u is following v and v has shared some news item n among their followers), predict whether the user u will also share news item n among their followers.*

We model the problem as a binary classification task and propose a set of features that takes into account user, news, and social network characteristics to predict better real and fake news sharing in social media. Our user-based features consider demographics, profile information, personality, emotions, user interest, and behavior, while our news-based features encode style, complexity, and psychological aspects of news headline and body [32]. Network-based features consider, instead, the user following network to measure tie strength and quantify opinion leadership. All these factors have never been combined into a unique predictive model or tested on a large scale before. We tested our approach on a Twitter dataset of 1,557 users built upon the FakeNewsNet [33] data and containing 7,661 user-news item sharing and not sharing instances of 169 fake news items. Our results show that our proposed set of features outperform the results of both baseline approaches, i.e., independent cascade and linear threshold models. Specifically, we show that our proposed features can predict fake news sharing with an AUROC of 97.34 and an average precision of 88.43 (vs. an AUROC of 67.70 and an average precision of 87.45

achieved by the best baseline). Among the proposed features, we observed that news-based features are more effective in predicting fake news sharing, followed by the user-based and network-based features.

B. Characterizing and Predicting Fake News Spreaders

Identifying fake news spreaders in social networks is one of the key aspects to mitigate misinformation spread effectively. Examples of strategies that could be implemented include assisting fake news spreaders with credibility indicators to lower their fake news sharing intent [34], and mitigation campaign, e.g., target the most influential real news spreader to maximize the spread of real news [33]. However, less is known about the characteristics of fake or real news spreaders.

Therefore, in our current work [35], we seek to understand the characteristics of fake news spreaders focusing on different attributes such as user writing style, emotions, demographics, personality, social media behavior, and network features. In particular, we leveraged these attributes to perform a comprehensive analysis on two different datasets, namely PolitiFact [33] and PAN [36] to investigate the patterns of user characteristics in social media in the presence of misinformation. We hypothesize that users likely to share fake news hold specific patterns based on these attributes, which are different from real news spreaders. To the best of our knowledge, some of the features we considered, such as user stress, needs, values, and tweeting behavior, have not been analyzed before. Furthermore, we investigate to what extent these features can be used to identify users who are likely to share fake news by addressing the problem as a binary classification task. Our experimental results show that using our proposed approach outperforms the results of baseline approaches with n-grams in both datasets. Specifically, we show that our proposed features can identify fake news spreaders with an average precision of 0.99 on the PolitiFact dataset (vs. 0.96 achieved by the best baseline) and 0.79 on the PAN dataset (vs. 0.78 achieved by the best baseline).

Furthermore, our analysis unveils some interesting characteristics of fake news spreaders across the two datasets considered. Specifically, we show that:

- The majority of users under 18 or over 40 may tend to share more fake than real news.
- Female users may tend to be more fake news spreaders than male users.
- The political orientation of a fake news spreader is more likely to coincide with the source's political bias of the majority of circulating fake news items.
- Fake news spreaders (1) have newer accounts, (2) spend, on average, less time between two consecutive tweets, and (3) tend to tweet more at night.
- Fake news spreaders tend to express more negative emotions and stress in their tweets than real news spreaders.
- Fake news spreaders are estimated to be more extroverted and less neurotic than real news spreaders.
- Emotions and personality features are strong predictors of fake news spreaders in all the considered datasets.

1) Predicting the Influence of Fake and Real News Spreaders: To further investigate the impact a fake news spreader can have on social media, we also address the problem of predicting the user influence as the number of retweets their latest news article tweet will receive. We address the problem as a multi-class classification problem and present a Random Forest classifier that categorizes the number of retweets a news tweet will receive into five ranges using user profile characteristics, emotion and sentiment analysis of tweets, and information about past tweets [37]. We use data from FakeNewsNet, containing a list of 43K known fake news spreaders and 135K real news spreaders, and the past 500 tweets of each user to build profiles of each user and compute the proposed features. This classifier results in a 0.931 and 0.853 weighted F1 score for real and fake news spreaders, respectively, performing better than other existing models, which resulted in a 0.928 weighted F1 score for real news spreaders and 0.832 F1 score for fake news spreaders at best. By comparing important features for predicting the influence of real and fake news spreaders, we show that an established account and the utilization of sources better characterize the influence of real news spreaders, while user interaction on Twitter is more important to determine the influence of fake news spreaders.

IV. Conclusions and Future Work

In this paper, we have presented current limitations of existing diffusion models for modeling misinformation diffusion in social media and described our current work on predicting whether an influenced user will share fake news and characterizing and predicting fake news spreaders and their influence. We have shown how our proposed methods that, inspired by the diffusion of innovation theory, consider several user- and news-based attributes were able to outperform existing approaches.

Future work will be devoted to studying real news diffusion as opposed to fake news, investigating similarities and differences of fake and real news diffusion models, and enhancing real and fake news diffusion models with a predicted time of when a user will share a news item they see on social media.

Acknowledgements

This research has been supported by the National Science Foundation under awards no. 1943370, 1950599, and 1820685. The author would like to thank her graduate students Anu Shrestha and Abishai Joy and her summer 2021 REU students Amy Zhang, Aaron Brookhouse, and Daniel Hammer for contributing to the presented research.

References

[1] A. KAPLAN, "Artificial intelligence, social media, and fake news: Is this the end of democracy?" *IN MEDIA & SOCIETY*, p. 149, 2020.

[2] F. Spezzano, A. Shrestha, J. A. Fails, and B. W. Stone, "That's fake news! reliability of news when provided title, image, source bias full article," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, no. CSCW1, Apr. 2021. [Online]. Available: https://doi.org/10.1145/3449183

[3] A. Shrestha and F. Spezzano, "Online misinformation: from the deceiver to the victim," in *ASONAM '19: International Conference on Advances in Social Networks Analysis and Mining, Vancouver, British Columbia, Canada, 27-30 August, 2019*. ACM, 2019, pp. 847–850.

[4] F. Spezzano, "Using service-learning in graduate curriculum to address teenagers' vulnerability to web misinformation," in *ITiCSE 2021: 26th ACM Conference on Innovation and Technology in Computer Science Education, Virtual Event, Germany, June 26 - July 1, 2021 - Working Group Reports*, C. Schulte, B. A. Becker, M. Divitini, and E. Barendsen, Eds. ACM, 2021, pp. 637–638.

[5] M. Gabielkov, A. Ramachandran, A. Chaintreau, and A. Legout, "Social clicks: What and who gets read on twitter?" in *Proceedings of the 2016 ACM SIGMETRICS international conference on measurement and modeling of computer science*, 2016, pp. 179–192.

[6] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of computational science*, vol. 2, no. 1, pp. 1–8, 2011.

[7] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy," in *Proceedings of the 22nd international conference on World Wide Web*, 2013, pp. 729–736.

[8] K. Starbird, J. Maddock, M. Orand, P. Achterman, and R. M. Mason, "Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing," *IConference 2014 Proceedings*, 2014.

[9] R. Kouzy, J. Abi Jaoude, A. Kraitem, M. B. El Alam, B. Karam, E. Adib, J. Zarka, C. Traboulsi, E. W. Akl, and K. Baddour, "Coronavirus goes viral: quantifying the covid-19 misinformation epidemic on twitter," *Cureus*, vol. 12, no. 3, 2020.

[10] L. Ma, C. S. Lee, and D. H. Goh, "Understanding news sharing in social media from the diffusion of innovations perspective," in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*. IEEE, 2013, pp. 1013–1020.

[11] S. Talwar, A. Dhir, P. Kaur, N. Zafar, and M. Alrasheedy, "Why do people share fake news? associations between the dark side of social media use and fake news sharing behavior," *Journal of Retailing and Consumer Services*, vol. 51, pp. 72–82, 2019.

[12] R. Zafarani, M. A. Abbasi, and H. Liu, *Social media mining: an introduction*. Cambridge University Press, 2014.

[13] E. M. Rogers, *Diffusion of innovations*. Simon and Schuster, 2010.

[14] X. Zhou and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Comput. Surv.*, vol. 53, no. 5, pp. 109:1–109:40, 2020.

[15] K. Shu, H. R. Bernard, and H. Liu, "Studying fake news via network analysis: detection and mitigation," in *Emerging research challenges and opportunities in computational social network analysis and mining*. Springer, 2019, pp. 43–65.

[16] N. Vo and K. Lee, "The rise of guardians: Fact-checking url recommendation to combat fake news," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018, pp. 275–284.

[17] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *ACM Computing Surveys (CSUR)*, vol. 51, no. 2, pp. 1–36, 2018.

[18] F. Jin, E. Dougherty, P. Saraf, Y. Cao, and N. Ramakrishnan, "Epidemiological modeling of news and rumors on twitter," in *Proceedings of the 7th workshop on social network mining and analysis*, 2013, pp. 1–9.

[19] G. Shrivastava, P. Kumar, R. P. Ojha, P. K. Srivastava, S. Mohan, and G. Srivastava, "Defensive modeling of fake news through online social networks," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 5, pp. 1159–1167, 2020.

[20] M. Tambuscio, G. Ruffo, A. Flammini, and F. Menczer, "Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks," in *Proceedings of the 24th international conference on World Wide Web*, 2015, pp. 977–982.

[21] T. Murayama, S. Wakamiya, E. Aramaki, and R. Kobayashi, "Modeling the spread of fake news on twitter," *PLOS ONE*, vol. 16, no. 4, pp. 1–16, 04 2021. [Online]. Available: https://doi.org/10.1371/journal.pone.0250419

[22] J.-H. Cho, S. Rager, J. O'Donovan, S. Adali, and B. D. Horne, "Uncertainty-based false information propagation in social networks," *ACM Transactions on Social Computing*, vol. 2, no. 2, pp. 1–34, 2019.

[23] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *SIGKDD*, 2003, pp. 137–146.

[24] C. S. Lee and L. Ma, "News sharing in social media: The effect of gratifications and prior experience," *Computers in human behavior*, vol. 28, no. 2, pp. 331–339, 2012.

[25] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

[26] K. Shu, X. Zhou, S. Wang, R. Zafarani, and H. Liu, "The role of user profiles for fake news detection," in *ASONAM '19: International Conference on Advances in Social Networks Analysis and Mining*. ACM, 2019, pp. 436–439.

[27] A. Guess, J. Nagler, and J. Tucker, "Less than you think: Prevalence and predictors of fake news dissemination on facebook," *Science advances*, vol. 5, no. 1, p. eaau4586, 2019.

[28] F. Rangel, A. Giachanou, B. Ghanem, and P. Rosso, "Overview of the 8th author profiling task at pan 2020: Profiling fake news spreaders on twitter," in *CLEF*, 2020.

[29] A. Shrestha, F. Spezzano, and A. Joy, "Detecting fake news spreaders in social networks via linguistic and personality features," in *CLEF*, 2020.

[30] A. Giachanou, E. A. Rissola, B. Ghanem, F. Crestani, and P. Rosso, "The role of personality and linguistic patterns in discriminating between fake news spreaders and fact checkers," in *Natural Language Processing and Information Systems: 25th International Conference on Applications of Natural Language to Information Systems, NLDB 2020*. Springer Nature, p. 181.

[31] A. Joy, A. Shrestha, and F. Spezzano, "Are you influenced? modeling the diffusion of fake news in social media," in *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2020, Virtual Event, Netherlands, November 8-11, 2021*.

[32] A. Shrestha and F. Spezzano, "Textual characteristics of news title and body to detect fake news: A reproducibility study," in *Advances in Information Retrieval - 43rd European Conference on IR Research, ECIR 2021, Virtual Event, March 28 - April 1, 2021, Proceedings, Part II*, ser. Lecture Notes in Computer Science, vol. 12657. Springer, 2021, pp. 120–133. [Online]. Available: https://doi.org/10.1007/978-3-030-72240-1_9

[33] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media," *arXiv preprint arXiv:1809.01286*, vol. 8, 2018.

[34] W. Yaqub, O. Kakhidze, M. L. Brockman, N. Memon, and S. Patil, "Effects of credibility indicators on social media news sharing intent," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20, 2020, p. 1–14.

[35] A. Shrestha and F. Spezzano, "Characterizing and predicting fake news spreaders in social networks," *International Journal of Data Science and Analytics (to appear)*, 2022.

[36] F. Rangel, A. Giachanou, B. Ghanem, and P. Rosso, "Overview of the 8th Author Profiling Task at PAN 2020: Profiling Fake News Spreaders on Twitter," in *CLEF 2020 Labs and Workshops, Notebook Papers*, L. Cappellato, C. Eickhoff, N. Ferro, and A. Névéol, Eds. CEUR Workshop Proceedings, Sep. 2020. [Online]. Available: CEUR-WS.org

[37] A. Zhang, A. Brookhouse, D. Hammer, F. Spezzano, and L. Babinkostova, "Predicting the influence of fake and real news spreaders (student abstract)," in *Thirty-Sixth AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, February 22- March 1, 2022*. AAAI Press, 2022, pp. –.