Fake News Research: Theories, Detection Strategies, and Open Problems

Reza Zafarani, Xinyi Zhou
Data Lab, EECS Department
Syracuse University
{reza,zhouxinyi}@data.syr.edu

Kai Shu, Huan Liu
Computer Science and Engineering
Arizona State University
{kai.shu,huan.liu}@asu.edu

ABSTRACT

Fake news has become a global phenomenon due its explosive growth, particularly on social media. The goal of this tutorial is to (1) clearly introduce the concept and characteristics of fake news and how it can be formally differentiated from other similar concepts such as mis-/dis-information, satire news, rumors, among others, which helps deepen the understanding of fake news; (2) provide a comprehensive review of fundamental theories across disciplines and illustrate how they can be used to conduct interdisciplinary fake news research, facilitating a concerted effort of experts in computer and information science, political science, journalism, social science, psychology and economics. Such concerted efforts can result in highly efficient and explainable fake news detection; (3) systematically present fake news detection strategies from four perspectives (i.e., *knowledge*, *style*, *propagation*, and *credibility*) and the ways that each perspective utilizes techniques developed in data/graph mining, machine learning, natural language processing, and information retrieval; and (4) detail open issues within current fake news studies to reveal great potential research opportunities, hoping to attract researchers within a broader area to work on fake news detection and further facilitate its development. The tutorial aims to promote a fair, healthy and safe online information and news dissemination ecosystem, hoping to attract more researchers, engineers and students with various interests to fake news research. Few prerequisite are required for KDD participants to attend.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing theory, concepts and paradigms**; • **Computing methodologies** → **Natural language processing**; • **Computing methodologies** → **Machine learning**; • **Security and privacy** → *Social aspects of security and privacy*; • **Applied computing** → *Sociology*; *Computer forensics.*

KEYWORDS

Fake news; fake news detection; news verification; false news; misinformation; disinformation; social media

ACM Reference Format:
Reza Zafarani, Xinyi Zhou and Kai Shu, Huan Liu. 2019. Fake News Research: Theories, Detection Strategies, and Open Problems. 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, 2 pages. https://doi.org/10.1145/3292500.3332287

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

ACM ISBN 978-1-4503-6201-6/19/08.
https://doi.org/10.1145/3292500.3332287

TUTORIAL OUTLINE

Fake news is now viewed as one of the greatest threats to democracy and journalism [17]. The reach of fake news was best highlighted during the critical months of the 2016 U.S. presidential election campaign, where top twenty frequently-discussed false election stories generated 8,711,000 shares, reactions, and comments on Facebook, ironically, larger than the total of 7,367,000 for the top twenty most-discussed election stories posted by top major news websites [11]. Our economies are not immune to fake news either, impacting stock markets and leading to massive trades. For example, fake news claiming that Barack Obama was injured in an explosion wiped out $130 billion in stock value [5, 14].

The generous benefits in fake news activities are one of the motivations for people to initiate and engage in such activities. Consider dozens of "well-known" teenagers in the Macedonian town of Veles who posted fake news for millions on social media and became wealthy by penny-per-click advertising during the 2016 U.S. presidential election [12]. Such stories attach greater importance to fake news detection and intervention as they provide an incentive for individuals to become the next "Macedonian teenagers" in the upcoming elections all around the world. With fake news detection research in its early stages, greater opportunities exist for such malicious individuals to create and spread fake news in the absence of a worry. On the other hand, it has been suggested that fake news is difficult to be recognized by the pubic, which leads to unintentional engagement in spreading fake news [17]; studies in social psychology and communications have demonstrated that human ability to detect deception is slightly better than chance, with a mean accuracy rate of 54% in over 100 experiments [8]. Such difficulty is also related to how individuals adjust (or correct) their judgments to fake news when it has already gained their trusts [7].

Facing such grim situation, this tutorial aims to (i) provide a clear understanding of fake news; (ii) attract researchers within general areas of data/graph mining, machine learning, Natural Language Processing (NLP), and Information Retrieval (IR) to conduct research on fake news and its detection and further facilitate its development; and (iii) encourage a collaborative effort of experts in computer and information science, political science, journalism, social science, psychology and economics to work on fake news detection, where such efforts can lead to fake news detection that is not only highly efficient, but more importantly, interpretable [9]. The tutorial contains the following four parts to achieve these goals:

I. Fake News and Related Concepts. We first present two definition of fake news in a broad and narrow way, which enables one to define fake news in terms of three general characteristics: (i) information authenticity, (ii) author intention, and (iii) whether the given information is in form of news. Such characteristics help differentiate fake news from the truth, as well as from several common related concepts, e.g., mis-/dis-information, satire news, and rumors. We will specify why fake news is defined in such ways,

what each characteristic indicates, and how it can be evaluated, quantified, or used to differentiate fake news from related concepts.

II. Fundamental Theories. Human vulnerability to fake news, which can bring in useful clues or further complicate fake news detection, has been a subject of interdisciplinary research [18]. For instance, achievements in forensic psychology such as *Undeutsch hypothesis* [13] have pointed out the style and quality differences between the truth and deceptive information. Similarly, interdisciplinary research has looked at why individuals spread fake information, considering that the borderline between malicious and normal users becomes unclear – normal people can also frequently and unintentionally participate in fake news activities, e.g., due to their social identity [1] or preexisting knowledge [4]. This tutorial conducts a comprehensive cross-disciplinary survey of literature on such theories. We review more than twenty well-known theories that can contribute to our understanding of fake news and participants in fake news activities [17]. We present and discuss the problems arising as explained by these theories, ranging from the patterns they can reveal, the qualitative and quantitative fake news studies one can conduct based on these studies, to the specific roles they can play in detecting fake news.

III. Detection Strategies. Detecting fake news is a complex and multidimensional task: it involves assessing multiple characteristics of news such as its authenticity, author intention, and its literary form. Furthermore, fake news is formed by multiple components (e.g., headline, body text, attached image(s)), and available information on fake news that can be utilized in predicting fake news sharply increases as it starts to disseminate online (e.g., feedback from users such as comments, its propagation paths on networks and its spreaders). Such components and information can be in the form of text, multimedia, network, etc., corresponding to various applicable techniques and usable resources.

To methodically and comprehensively present the ways to detect fake news, in this tutorial, we will specify how fake news detection can be conducted respectively from four perspectives (i.e., *knowledge*, *style*, *propagation* and *credibility*) - their corresponding general strategies, targeting fake news characteristic that can be evaluated, components and information that can be utilized, applicable techniques, and some typical approaches.

Generally speaking, fake news detection from a knowledge perspective is a "comparison" between the relational textual knowledge extracted from to-be-verified news articles and that of knowledge graphs representing facts or ground truth [2, 6]. The construction of knowledge graphs is an active research area within IR. Such "comparison" is often reduced to a link prediction (or knowledge inference) task, which directly evaluates news authenticity. Style-based fake news detection aims to capture the differences in writing styles between fake and true news, which often relies on NLP techniques and is conducted within a machine learning framework. News style can be extracted from the text [16], images [15], and/or videos within to-be-verified content, enabling one to indirectly evaluate the intention of the creator of news articles. Propagation-based and credibility-based fake news detection both further exploit information provided in news propagation on social media, where the former mainly relies on news cascades or self-defined graphs [14], while the latter emphasizes on exploring the credibility relationships between news articles and entities such as clickbait, publishers, spreaders, comments, etc. [3, 10]. Hence, research tasks involved can be correlated to clickbait detection, opinion spam detection, and the like. Here, graph optimization algorithms often play an important role to solve the target problems.

IV. Open Issues. In the final section of the tutorial, we will present the challenges and open issues that are important but have not been addressed (or thoroughly addressed) in current studies. Such challenges and open issues are three-fold: (i) challenges brought from news characteristics, e.g., the timeliness of news articles demands real-time knowledge graphs that can assure knowledge timeliness; (ii) open issues attached to model explainability; and (iii) open issues attached to model performance, e.g., the completeness of knowledge graphs and cross-domain generalization of style-based approaches. Five tasks, namely *fake news early detection*, *check-worthy content identification*, *cross-domain/topic/language study of fake news*, *representation learning for fake news detection*, and *fake news intervention*) will be thus highlighted, with discussions on why these tasks are crucial and potential ways to address each task.

TUTORIAL WEBSITE

The tutorial website is https://www.fake-news-tutorial.com/, where the slides, related papers, datasets and tools have been uploaded and will be timely updated.

REFERENCES

[1] Blake E Ashforth and Fred Mael. 1989. Social identity theory and the organization. *Academy of management review* 14, 1 (1989), 20–39.

[2] Xin Dong, Evgeniy Gabrilovich, Geremy Heitz, Wilko Horn, Ni Lao, Kevin Murphy, Thomas Strohmann, Shaohua Sun, and Wei Zhang. 2014. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 601–610.

[3] Zhiwei Jin, Juan Cao, Yongdong Zhang, and Jiebo Luo. 2016. News Verification by Exploiting Conflicting Social Viewpoints in Microblogs.. In *AAAI*. 2972–2978.

[4] Raymond S Nickerson. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology* 2, 2 (1998), 175.

[5] K Rapoza. 2017. Can 'fake news' impact the stock market?

[6] Xiang Ren, Nanyun Peng, and William Yang Wang. 2018. Scalable Construction and Reasoning of Massive Knowledge Bases. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorial Abstracts*. 10–16.

[7] Arne Roets et al. 2017. 'Fake news': Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions. *Intelligence* 65 (2017), 107–110.

[8] Victoria L Rubin. 2010. On deception and deception detection: Content analysis of computer-mediated stated beliefs. *Proceedings of the Association for Information Science and Technology* 47, 1 (2010), 1–10.

[9] Kai Shu, Limeng Cui, Suhang Wang, Dongwon Lee, and Huan Liu. 2019. dEFEND: Explainable Fake News Detection. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM.

[10] Kai Shu, Suhang Wang, and Huan Liu. 2019. Beyond News Contents: The Role of Social Context for Fake News Detection. In *WSDM*. https://doi.org/10.1145/3289600.3291382

[11] Craig Silverman. 2016. This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed News* 16 (2016).

[12] Alexander Smith and Vladimir Banic. 2016. Fake News: How a partying Macedonian teen earns thousands publishing lies. *NBC News* 9 (2016).

[13] Udo Undeutsch. 1967. Beurteilung der glaubhaftigkeit von aussagen. *Handbuch der psychologie* 11 (1967), 26–181.

[14] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151.

[15] Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun, Kishlay Jha, Lu Su, and Jing Gao. 2018. EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 849–857.

[16] Xinyi Zhou, Atishay Jain, Vir V Phoha, and Reza Zafarani. 2019. Fake News Early Detection: A Theory-driven Model. *arXiv preprint arXiv:1904.11679* (2019).

[17] Xinyi Zhou and Reza Zafarani. 2018. Fake News: A Survey of Research, Detection Methods, and Opportunities. *arXiv preprint arXiv:1812.00315* (2018).

[18] Xinyi Zhou and Reza Zafarani. 2019. Fake News Detection: An Interdisciplinary Research. In *Companion of The Web Conference. https://doi.org/10.1145/3308560.3316476.*