The Battlefront of Combating Misinformation and Coping with Media Bias

Yi R. Fung
University of Illinois Urbana-Champaign
USA
yifung2@illinois.edu

Kung-Hsiang Huang
University of Illinois Urbana-Champaign
USA
khhuang3@illinois.edu

Preslav Nakov
Mohamed bin Zayed University of Artificial Intelligence
UAE
preslav.nakov@mbzuai.ac.ae

Heng Ji
University of Illinois Urbana-Champaign
USA
hengji@illinois.edu

ABSTRACT

Misinformation is a pressing issue in modern society. It arouses a mixture of anger, distrust, confusion, and anxiety that cause damage on our daily life judgments and public policy decisions. While recent studies have explored various fake news detection and media bias detection techniques in attempts to tackle the problem, there remain many ongoing challenges yet to be addressed, as can be witnessed from the plethora of untrue and harmful content present during the COVID-19 pandemic, which gave rise to the first social-media infodemic, and the international crises of late. In this tutorial, we provide researchers and practitioners with a systematic overview of the frontier in fighting misinformation. Specifically, we dive into the important research questions of how to (i) develop a robust fake news detection system that not only fact-checks information pieces provable by background knowledge, but also reason about the consistency and the reliability of subtle details about emerging events; (ii) uncover the bias and the agenda of news sources to better characterize misinformation; as well as (iii) correct false information and mitigate news biases, while allowing diverse opinions to be expressed. Participants will learn about recent trends, representative deep neural network language and multimedia models, ready-to-use resources, remaining challenges, future research directions, and exciting opportunities to help make the world a better place, with safer and more harmonic information sharing.

ACM Reference Format:
Yi R. Fung, Kung-Hsiang Huang, Preslav Nakov, and Heng Ji. 2022. The Battlefront of Combating Misinformation and Coping with Media Bias. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD ’22). ACM, New York, NY, USA, 2 pages. https://doi.org/XXXXXXX.XXXXXXX

1 PRESENTER BIOGRAPHIES

Yi R. Fung is a third-year CS Ph.D. candidate at the University of Illinois Urbana-Champaign (UIUC), with research interests in NLP, multimedia reasoning, and computation for social good. Her recent research includes fake news detection and cross-culture understanding. Yi is a recipient of the NAAACL’21 Best Demo Paper and a UIUC Lauslen and Andrew fellowship. She gave an invited talk at the Harvard Medical School Bioinformatics Seminar.

Kung-Hsiang Huang is a second-year Ph.D. student at the Computer Science Department of UIUC. His research focuses on fact-checking and fake news detection. Prior to joining UIUC, he obtained his B.Eng. in Computer Science from the Hong Kong University of Science and Technology, and his M.Sc. in Computer Science from USC. He is also a co-founder of an AI startup, Rosetta.ai.

Preslav Nakov is a Professor at Mohamed bin Zayed University of Artificial Intelligence. He is President of ACL SIGLEX, Secretary of ACL SIGSLAV, advisory board member of EACL, and editorial board member of Computational Linguistics, TACL, CS&L, IEEE TAC, NLE, AI Communications, and Frontiers in AI. His research on fake news has been featured by over 100 news outlets, including Forbes, Boston Globe, Aljazeera, MIT Technology Review, The Register, WIRED, and Engadget, among others.

Heng Ji is a Professor at the Computer Science Department of the University of Illinois Urbana-Champaign, and an Amazon Scholar. Her research interests focus on NLP, especially Multimedia Multilingual Information Extraction, Knowledge Base Population and Knowledge-driven Generation. She was selected as “Young Scientist” and a member of the Global Future Council on the Future of Computing by the World Economic Forum. She received “AI’s 10 to Watch” Award, NSF CAREER award, Google Research Award, IBM Watson Faculty Award, Bosch Research Award, Amazon AWS Award, ACL2020 and NAACL2021 Best Demo Paper Award.

2 TUTORIAL OUTLINE

2.1 Background and Motivation [20min]
We begin with a selection of real-world examples of fake news and their harmful impact on society, followed by a pedagogical exercise of how humans tend to approach the problem of fake news detection, characterization, and correction. We point out the conceptual distinctions between different types of fake news.
2.2 Fake News Detection [60min]

We discuss detection based on stylistic [6, 19], fact-checking [13, 18, 22], semantic consistency [2, 14], and propagation patterns [1, 26], and their advantages and limitations [21, 27]. We then cover approaches [11, 16] that leverage a combination of these elements for greater representation power and robustness. Importantly, we discuss work that explores the diachronic bias of fake news detection and portability across datasets in different time and language settings [12, 17]. Moreover, we review generating fake news that better aligns with the key topic and the facts [11, 25], e.g., by applying conditioned fake news generation to construct silver-standard data annotations for finer-grained fake news detection [11].

2.3 Fake News Characterization [30min]

We next address some fundamental questions of characterization based on underlying source biases, reporting agenda, propaganda techniques, and target audience [5]. We introduce approaches to detect political and socio-cultural biases in news articles [3, 9, 10]. Next, we discuss the recent EMU benchmark that requires models to answer open-ended questions capturing the intent and the implications of a media edit [7]. We cover methodologies for identifying specific propaganda techniques, e.g., loaded language, appeal to fear, smears, glittering generalities, whataboutism, etc. [8]. We also discuss recent exploration in predicting the intended target of harmful memes, e.g., person, organization, community, or society [20, 23].

2.4 Corrective Actions [30min]

Finally, we cover research on explaining why a given piece of news is fake through the leverage of reader comments [24], strategies for placing corrective explanations, based on user studies [4], and mitigating media bias through neutral article generation [15].

ACKNOWLEDGEMENTS

This work is supported in part by U.S. DARPA SemaFor Program No. HR001120C0123, AIDA Program No. FA8750-18-2-0014, and KAIROS Program No. FA8750-19-2-1004. The views and the conclusions contained herein are those of the authors and should not be interpreted as representing the official policies of DARPA or the U.S. Government. The U.S. Government is authorized to reproduce and to distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

REFERENCES

[1] S. Aneja, C. Bregler, and M. Niehner. Cosmos: Catching out-of-context misinformation with self-supervised learning. arXiv preprint arXiv:2106.06728, 2021.
[2] Anonymous. Cross-document misinformation detection based on event graph reasoning. Open Review preprint, 2022.
[3] R. Baly, G. Da San Martino, J. Glass, and P. Nakov. We can detect your bias: Predicting the political ideology of news articles. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP ’20, pages 4982–4991, 2020.
[4] N. M. Brashear, G. Pennycook, A. J. Berinsky, and D. G. Rand. Timing matters when correcting fake news. Proceedings of the National Academy of Sciences, 118(5), 2021.
[5] T. Buchanan. Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. PLoS One, 15(10):e0236666, 2020.
[6] L. Choshen, D. Eldad, D. Hershcovich, E. Sulem, and O. Abend. The language of legal and illegal activity on the Darknet. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, ACL ’19, pages 4271–4279, Florence, Italy, 2019.
[7] J. Da, M. Forbes, R. Zellers, A. Zheng, J. D. Hwang, A. Bosselut, and Y. Choi. Edited media understanding frames: Reasoning about the intent and implications of visual misinformation. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL-IJCNLP ’21, pages 2026–2039, 2021.
[8] D. Dimitrov, B. Bin Ali, S. Shaar, F. Alam, F. Silvestri, H. Firooz, P. Nakov, and G. Da San Martino. Detecting propaganda techniques in memes. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL-IJCNLP ’21, pages 6605–6617, 2021.
[9] L. Fan, M. White, E. Sharma, R. Su, P. K. Choubey, R. Huang, and L. Wang. In plain sight: Media bias through the lens of factual reporting. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP ’19, pages 6343–6349, Hong Kong, China, 2019.
[10] M. Forbes, J. D. Hwang, V. Shwartz, M. Sap, and Y. Choi. Social chemistry 101: Learning to reason about social and moral norms. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP ’20, pages 653–670, 2020.
[11] Y. Fung, C. Thomas, R. G. Reddy, S. Polisetti, H. Ji, S.-F. Chang, K. McKeown, M. Bansal, and A. Sil. InfoSurgeon: Cross-media fine-grained information consistency checking for fake news detection. In Proceedings of the Annual Meeting of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing, ACL-IJCNLP ’21, pages 1683–1698, 2021.
[12] F. Gereme, W. Zhu, T. Ayali, and A. Alemu. Combating fake news in “low-resource” languages: Andaric fake news detection accompanied by resource crafting. Information, 12(3):29, 2021.
[13] L. Hu, T. Yang, L. Zhang, W. Zhong, D. Tang, C. Shi, N. Duan, and M. Zhou. Compare to the knowledge: Graph neural fake news detection with external knowledge. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL-IJCNLP ’21, pages 754–763, 2021.
[14] H. Karimi and J. Tang. Learning hierarchical discourse-level structure for fake news detection. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT ’19, pages 3432–3442, Minneapolis, MN, USA, 2019.
[15] N. Lee, Y. Bang, A. Madotto, and P. Fung. Mitigating media bias through neutral article generation. arXiv preprint arXiv:2104.00336, 2021.
[16] N. Lee, B. Z. Li, S. Wang, P. Fung, H. Ma, W.-Y. Yih, and B. Khabsa. On unifying misinformation detection. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT ’21, pages 5479–5485, 2021.
[17] T. Murayama, S. Wakamiya, and E. Aramaki. Mitigation of diachronic bias in fake news detection dataset. In Proceedings of the Seventh Workshop on Noisy User-generated Text, W-NUT ’21, pages 182–188, 2021.
[18] V.-H. Nguyen, K. Sugiyama, P. Nakov, and M.-Y. Kan. FANG: Leveraging social context for fake news detection using graph representation. Commun. ACM, 65(4):124–132, mar 2022.
[19] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea. Automatic detection of fake news. In Proceedings of the 27th International Conference on Computational Linguistics, COLING ’18, pages 3391–3401, Santa Fe, NM, USA, 2018.
[20] S. Pramanick, S. Sharma, D. Dimitrov, M. S. Akhtar, P. Nakov, and T. Chakraborty. MOMENTA: A multimodal framework for detecting harmful memes and their targets. In Findings of the Association for Computational Linguistics: EMNLP 2021, pages 4439–4455, Punta Cana, Dominican Republic, 2021.
[21] T. Schuster, R. Schuster, D. J. Shah, and R. Barzilay. The limitations of stylometry for detecting machine-generated fake news. Computational Linguistics, 46(2):499–510, 2020.
[22] S. Shaar, N. Babulkov, G. Da San Martino, and P. Nakov. That is a known lie: Detecting previously fact-checked claims. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL ’20, pages 3607–3618, 2020.
[23] S. Sharma, M. S. Akhtar, P. Nakov, and T. Chakraborty. DISARM: Detecting the victims targeted by harmful memes. In Findings of NAACL 2022, Seattle, WA, USA, 2022.
[24] K. Shu, L. Cui, S. Wang, D. Lee, and H. Liu. dEFEND: Explainable fake news detection. In Proceedings of the 25th International Conference on Knowledge Discovery & Data Mining, KDD ’19, pages 395–405, Anchorage, AK, USA, 2019.
[25] K. Shu, Y. Li, K. Dung, and H. Liu. Fact-enhanced synthetic news generation. Proceedings of the AAAI Conference on Artificial Intelligence, 35(15):13802–13807, 2021.
[26] R. Tan, B. Plummer, and K. Saenko. Detecting cross-modal inconsistency to defend against neural fake news. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP ’20, pages 2081–2106, 2020.
[27] K. Zhou, C. Shu, B. Li, and J. H. Lau. Early rumour detection. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT ’19, pages 1614–1623, Minneapolis, MN, USA, 2019.