Published in : Zimmermann, A., Howlett, R.J., Jain, L.C. (eds) Human Centred Intelligent Systems. Smart Innovation, Systems and Technologies, vol 310. Springer, Singapore. https://doi.org/10.1007/978-981-19-3455-1_1
which should be cited to refer to this work.

A Collaborative Intelligence Approach to Fighting COVID-19 False News: A Chinese case

Zhan Liu[1][0000−0003−3367−3204], Shaban Shabani[2][0000−0003−4710−6091], Xianyun Yu[3][0000−0002−2336−4240], Maria Sokhn[2][0000−0001−7586−0564], and Nicole Glassey Balet[1][0000−0003−3268−8375]

[1] University of Applied Sciences and Arts Western Switzerland (HES-SO Valais-Wallis), Switzerland
{zhan.liu; nicole.glassey}@hevs.ch
[2] University of Applied Sciences Western Switzerland (HES-SO Neuchâtel), Switzerland
{shaban.shabani; maria.sokhn}@he-arc.ch
[3] Hangzhou Dianzi University, China
xyyu@hdu.edu.cn

Abstract. The rapid outbreak of COVID-19 has heightened interest in news about the pandemic. In addition to obtaining real-time developments about COVID-19, people have learned about prevention methods through the news media. Ironically, false COVID-19 news has spread faster than the virus, posing an additional health threat with advice being as dangerous as infection. In this study, we developed a Chinese news article dataset on COVID-19 misinformation, which contained 1266 verified articles from 118 Chinese digital newspaper platforms from January 2020 to January 2021. This dataset uses machine learning methods to detect false news in the Chinese language. Because automated classification methods, combined with human computation-based approaches, are effective for combating digital misinformation, we applied and evaluated a collaborative intelligence approach that leverages human fact-checking skills with feedback on news stories using four criteria: source, author, message, and spelling. The results show that reliable human feedback can help detect false news with high accuracy.

Keywords: collaborative intelligence · COVID-19 · false news · news articles · logistic regression · human-centred approach.

1 Introduction

Although misinformation is everywhere and much of it is harmless, the spread of COVID-19 fake news generates unique problems and dangers to public health. Because of this, fighting fake coronavirus news has become a significant challenge for both researchers and practitioners. In the words of T.A. Ghebreyesus [1], the general director of the World Health Organization, fake news on Weibo and WeChat, two microblogging websites, can spread extraordinarily fast, and most fake news stories with health advice range from useless (and relatively harmless) to extremely dangerous (and sometimes fatal). Additionally, many conspiracy stories minimize the seriousness of the virus and heighten the risk of infection. To counter this, and in light of the high level of user distrust toward fact-checkers, readers should be empowered to recognize fake news and understand its consequences. As a result, we identified a user-friendly tool to raise awareness of and identify COVID-19 related fake news, and help prevent its dissemination.

The first case of COVID-19 was identified in China in December 2019 [22] and related information from China is now often referenced by news reporters and scholars. However, the COVID-19 pandemic has given rise to significant misinformation and conspiracy theories about the pandemic's scale, origin, disease prevention, diagnosis, and treatment. This misinformation is widely circulated not only on social media but also on digital news media. While many reputable news providers attempt to mitigate the effects of fake news, it remains difficult to provide reliable COVID-related information from verified sources. For this reason, some news outlets have abandoned the breaking news approach [9][24], turning to fact-checking services such as WikiTribune[1], PolitiFact[2], Snopes[3], and The Washington Post, among others. These manually fact-check news items and verify articles, books, government agency statistics, photographs, and recorded interviews. Although the services often rank website pages or write fact-check evaluations, many readers are unable to distinguish by themselves the authenticity of the news due to a lack of knowledge and trust in the media.

In recent years, however, several methods have emerged to help readers detect fake news. Liu et al. [10] and Tschiatschek et al. [23], for example, used technical methods to detect fake news: the former applied machine learning for automated classification, and the latter leveraged crowd signals to discover fake news on social media. Humprecht [8], after summarizing the characteristics of fake news and analyzing its different effects in different countries, found that online disinformation is not only a technology-driven phenomenon but is also shaped by national information environments. While these and other authors have applied human-machine collaboration methods in their studies [19, 3, 14, 2, 22, 20], which play an important role in improving the accuracy of fake news detection, bringing together IT experts, universities, and media to devise technologies that can help journalists to find fake online claims and for readers to verify uncertain sources. However, little research has thus far focused on COVID-19 related fake news detection, and there is no relevant dataset in Chinese for machine learning training. Thus, because few COVID-19 datasets are used in machine learning classification, the accuracy of the results remains unsatisfactory.

In the present study, we built a fact-checked Chinese COVID-19 textual dataset containing 1266 news articles from 118 Chinese digital newspaper platforms. This dataset provides not only a precondition for analyzing the content patterns in COVID-19 fake news but also enables further research into detect-

[1] https://wt.social/
[2] https://www.politifact.com/
[3] https://www.snopes.com/fact-check/

ing misinformation. Moreover, we propose a collaborative intelligence approach SAMS, which uses a linear classification model [13] to classify news articles into true and false. Our approach leverages the fact-checking skills of humans to detect COVID-19 related fake news articles, which is integrated into the machine learning process to improve the accuracy of the classification process.

The remainder of this paper is organized as follows. Section 2 presents the literature review. Section 3 presents our method, including how we collect data, and framework, and Section 4 introduces the implementation and preliminary evaluation results. Finally, we conclude with a summary of the current work and offer suggestions for future research.

2 Related Work

Misinformation about COVID-19, like the pandemic itself, is now global. Since the first cases were confirmed in China in 2019, the voices of conspiracy theorists and alternative media have amplified, exploiting both the crisis and social media platforms to spread rumors. Recently, a set of social media datasets about COVID-19 emerged. Cui et Lee [5] provided a CoAID dataset that consists of true and false news about COVID-19 from diverse sources. This dataset covered 3252 news articles on websites and 851 posts on social network platforms. Furthermore, Chen et al. [4] published their coronavirus dataset from Twitter, beginning in January 2020 when they started collecting the data. Today, there are more than 1.8 billion tweets on their GitHub repository. Similarly, Patwa et al. [16] released a manually annotated dataset of 10700 social media posts of real and fake news on COVID-19 from Facebook, Twitter, and Instagram. While most of these studies focused on English-language-based datasets, only a few social media COVID-19 datasets currently exist in Chinese. One exception is Luo et al. [11], who published a Chinese infodemic dataset—"infodemic 2019"—by collecting widely spread Chinese infodemic from Sina Weibo microblogs during the COVID-19 outbreak. This dataset contains 1055 records, of which 409 entries were 'questionable', 276 'false', and 335 'true' from 21 January to 10 April 2020. Checked [25] is another Chinese dataset on COVID-19 misinformation, and contains 2104 verified microblogs related to COVID-19 from Weibo between December 2019 and August 2020. In this dataset, additional information, such as reposts, comments, and likes were included. Nevertheless, few Chinese COVID-19 related news article datasets have been constructed to support the fact-checking research, a gap this study attempts to fill.

To improve the accuracy of fact-checking, especially in identifying true and false news reports, an increasing number of studies have integrated human-based methods with the development of machine learning classification [14, 19, 15, 17]. Collaborative intelligence (CI), as a shared human group intelligence approach, relies on group cooperation and efforts to achieve consensus decision-making [12], a method now widely used in many disciplines. Gregg [7] demonstrated the requirement for designing CI applications, in which he developed the "DDtrac" application for children with special needs. This application was intended to support decision-making in special education and therapy, allowing data collection and providing tools for data analysis to understand a child's progress and adjust the teaching where necessary. Elia et al. [6] proposed a framework for the digital entrepreneurship ecosystem using a collective intelligence approach. The authors presented four dimensions associated with digital actors (who), digital activities (what), digital motivations (why), and digital organization (how). Moreover, they suggested how the entrepreneurial process can take advantage of the platform-based innovation ecosystem and collective intelligence approach. To understand how the healthcare ecosystem is challenged by COVID-19, a recent study by Secundo et al. [18] proposed a framework using digital technologies based on the collective intelligence principle. Their analysis introduced policy implications based on a unique decision support system to allocate a limited set of IoT devices to a larger group of patients, balancing alternative needs to improve the condition of severely ill patients while maximizing the efficiency of device use. The collaborative intelligence approach, however, has not been used to detect COVID-19 related news articles.

3 Methods and Framework

In this section, we explain how we collected the data and information in the dataset. We then introduce SAMS, a collaborative intelligence approach for fake news detection.

3.1 Data Collection

We defined a group of guidelines in the data collection process to obtain the COVID-19 related Chinese news articles with the ground truth labels. Our guidelines are described below.

- Label data. To collect reliable COVID-19 news, we rely on official and reputable news reports, including those from the Chinese Center for Disease Control and Prevention (CCDC), as well as national and local official online media. News articles from these sources are highly credible. On the other hand, when collecting fake news, we need to ensure that each falsehood has been defined by experts and that their detailed evaluation of the reports is published. Some existing platforms such as Jiaozhen from Tencent and Dajia from Weibo have been designed for news verification and can be used as channels for our fake news collection.
- Diversify sources and topics. To avoid common writing style issues, we included many news articles from different online media. In addition, to avoid classification bias, we balanced the number of true and fake news reports, to develop a database covering various topics, including health, economics, politics, sports, culture, and technology.

By following the data collection guidelines, we filtered true and false news from 126 Chinese digital media platforms between January 2020 and January

2021. Our final sample included 810 true and 422 false stories. Table 1 shows the descriptive statistics of the article length. The average length of true and false news articles is 333 and 184 words, respectively. Table 1 presents an overview of the dataset and details concerning article length.

Table 1. Dataset overview

Category	Samples	Title Length	Snippet Length	Total Length
True	810	18.5	164	184
False	422	19.8	312	333

To reach a broader distribution of the sources and topics, we collected up to 15 articles from each newspaper platform, which covered at least three different topics. Each news article contains the following components:

– ID: unique identity in the format of digital number
– Topic: subject of the article. If possible, the location of the news is attached
– Publisher: name of the publisher
– Title: title of the article
– Snippet: summary of the article
– Full text: full-textual information of the news article
– URL: address of the article. For fake news stories, the URL refers to the content as shown in Figure 1, which includes the false information, analysis report, the expert's name, and the location where the result was published
– Publish date: the time of publication on the newspaper platform (yyyy-mm-dd)
– Source: where the article comes from. This information might be missed in the fake news dataset because of removal from the original website
– Label: real or fake labeling for the news article

3.2 SAMS - collaborative intelligence approach

Although machine learning is a widely used tool for fake news detection, research of mixed-initiative fact-checking [14, 19] suggests that machine learning alone is not as accurate as integrating the human element into the fact-checking process. To do this, we proposed a collaborative intelligence approach-based system with the linear machine learning classification, to detect COVID-related fake news articles in Chinese. In our approach, we followed and adapted the guidelines of SAMS [19], in which four important features are identified to perform fact-checking: source, author, message, and spelling. Automatically answering questions about these four features is non-viable, and humans can do better by performing fact-checking skills such as searching for facts on trustworthy data resources. We define a process with tips and tricks to easily answer each of the questions for SAMS.

Fig. 1. An example of a fake news article with an expert's evaluation on the Jiaozhen platform

- **Source** - taking a critical look at the source, both data, and metadata, is the first step. The goal is to understand if the Chinese news content has sources and if they are reliable.
- **Author** - identifying whether there is an author for the news item. In principle, real and serious news articles usually have an author. It is necessary to check for related publications by the same author.
- **Message** - checking whether the message is clear, balanced, and fair. It should not contain unsupported or outrageous statements.
- **Spelling** - verifying grammar mistakes, such as repeated spelling mistakes, poor grammar, incorrect punctuation. Reputable news articles are carefully proofread before publication.

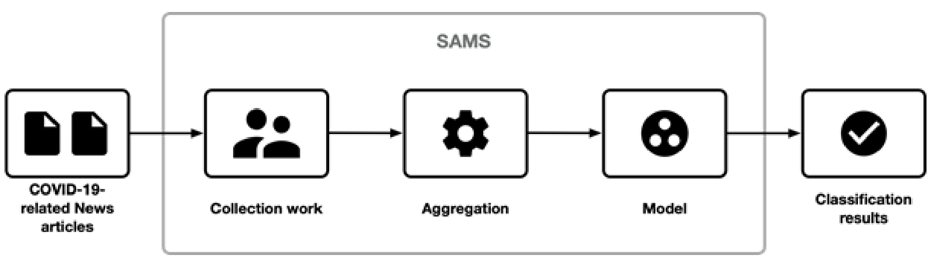

Fig. 2. Proposed collaborative intelligence framework following the SAMS approach[19]

Figure 2 illustrates the method's framework—the collaborative intelligence approach-based component, which aggregates the inputs from multiple users to infer the true answers related to SAMS questions. The output of the users' answers is a vector of four continuous values. The output of SAMS is then aggregated into a feature vector. Finally, the generated features are used for training and evaluating the machine learning models.

4 Implementation and Evaluation

While machine learning approaches are now widely used to detect false news on social media [21], because of the difficulty of classifying using algorithmic methods only, other approaches rely on crowdsourcing [23] or hybrid human-machine [20, 19] that leverage cognitive skills of online users.

In this study, we focus on evaluating the importance of human-provided feedback via the SAMS [19] features. Initially, two annotators provided answers to SAMS questions as described in Section 3.2. The answer was rated on a scale of 0 to 1, where 1 is 'completely agree with the criteria', and 0 indicates the opposite, i.e., not applicable. The scores from two annotators were aggregated by averaging the values and generating a vector with four features. Because of the low number of features, we use a logistic classification model that models the probability of the news article belonging to one of two categories, true or false. We split the dataset into training and testing, assigning 80% of the data for training the model and 20% for evaluating the performance of the classification model. Since the two dataset classes are not equally represented (810 true, 422 false new articles), the training (true: 648 samples, false: 337 samples) and testing (true: 162 samples, false: 85 samples) sets were sampled to preserve the dataset distribution.

Table 2. Evaluation results

	Reference	
Prediction	Real	False
Real	162	3
False	0	82

Accuracy	Sensitivity	Specificity
0.98	1.0	0.96

The results are presented in Table 2. The classification accuracy of the model is 98%, where the positive class is considered the *real* news articles. The sensitivity is 100%, meaning the model correctly classifies all the true news (162/162), whereas the specificity is 96%: 82 out of 85 false articles were correctly classified as false and 3 were misclassified as true. The high accuracy results can be justified because the SAMS scores were provided by annotators who are well-trained and understand the topic of false news.

5 Conclusion

In this study, we built a Chinese news article dataset of COVID-19 misinformation, containing 1266 verified news articles from 118 Chinese digital newspaper platforms from January 2020 to January 2021. Each news article contains ground truth labels, which could serve as a contribution to the COVID-19 false news research community. Moreover, we propose a collaborative intelligence approach, focusing on human-centered feedback to examine the questions from source, message, author, and spelling in the news article. Our preliminary results show that the four SAMS features are the most important in the classification model and are an effective (highly accurate) mechanism for combating digital misinformation. Our future research will focus on false news detection using a data mining approach, combining statistical and sentiment-based features automatically extracted from the text of the Chinese news articles.

Acknowledgments

The work described in this paper is supported by the Leading House for the Bilateral Science and Technology Cooperation Programme with Asia (ETH Zurich) under grant number COV_11_062020, and by the University of Applied Sciences and Arts Western Switzerland (HES-SO Valais-Wallis) under grant number 104353.

References

1. Munich security conference, https://www.who.int/director-general/speeches/detail/munich-security-conference
2. Albahar, M.: A hybrid model for fake news detection: Leveraging news content and user comments in fake news. IET Information Security **15**(2), 169–177 (2021). https://doi.org/10.1049/ise2.12021, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1049/ise2.12021
3. Bailer, W., Thallinger, G., Backfried, G., Thomas-Aniola, D.: Challenges for Automatic Detection of Fake News Related to Migration : Invited paper. In: 2021 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA). pp. 133–138 (May 2021). https://doi.org/10.1109/CogSIMA51574.2021.9475929, iSSN: 2379-1675
4. Chen, E., Lerman, K., Ferrara, E.: Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. JMIR Public Health and Surveillance **6**(2), e19273 (May 2020). https://doi.org/10.2196/19273
5. Cui, L., Lee, D.: Coaid: Covid-19 healthcare misinformation dataset. ArXiv **abs/2006.00885** (2020)
6. Elia, G., Margherita, A., Passiante, G.: Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. Technological Forecasting and Social Change **150**, 119791 (Jan 2020). https://doi.org/10.1016/j.techfore.2019.119791
7. Gregg, D.G.: Designing for collective intelligence. Communications of the ACM **53**(4), 134–138 (Apr 2010). https://doi.org/10.1145/1721654.1721691
8. Humprecht, E.: Where 'fake news' flourishes: a comparison across four Western democracies. Information, Communication & Society **22**(13), 1973–1988 (May 2018). https://doi.org/10.1080/1369118X.2018.1474241, https://www.tandfonline.com/doi/full/10.1080/1369118X.2018.1474241

9. Lewis, J., Cushion, S.: The Thirst to Be First. Journalism Practice **3**(3) (Aug 2009). https://doi.org/10.1080/17512780902798737, publisher: Routledge
10. Liu, Z., Shabani, S., Balet, N.G., Sokhn, M.: Detection of Satiric News on Social Media: Analysis of the Phenomenon with a French Dataset. In: 2019 28th International Conference on Computer Communication and Networks (ICCCN). pp. 1–6 (Jul 2019). https://doi.org/10.1109/ICCCN.2019.8847041, iSSN: 2637-9430
11. Luo, J., Xue, R., Hu, J., El Baz, D.: Combating the Infodemic: A Chinese Infodemic Dataset for Misinformation Identification. Healthcare **9**(9), 1094 (Sep 2021). https://doi.org/10.3390/healthcare9091094, number: 9 Publisher: Multidisciplinary Digital Publishing Institute
12. Malone, T.W., Laubacher, R., Dellarocas, C.: Harnessing Crowds: Mapping the Genome of Collective Intelligence. SSRN Scholarly Paper ID 1381502, Social Science Research Network, Rochester, NY (Feb 2009). https://doi.org/10.2139/ssrn.1381502
13. Mladenić, D., Brank, J., Grobelnik, M., Milic-Frayling, N.: Feature selection using linear classifier weights: interaction with classification models. In: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval. pp. 234–241. SIGIR '04, Association for Computing Machinery, New York, NY, USA (Jul 2004). https://doi.org/10.1145/1008992.1009034
14. Nasir, J.A., Khan, O.S., Varlamis, I.: Fake news detection: A hybrid CNN-RNN based deep learning approach. International Journal of Information Management Data Insights **1**(1), 100007 (Apr 2021). https://doi.org/10.1016/j.jjimei.2020.100007
15. Nguyen, A.T., Kharosekar, A., Krishnan, S., Krishnan, S., Tate, E., Wallace, B.C., Lease, M.: Believe it or not: Designing a Human-AI Partnership for Mixed-Initiative Fact-Checking. In: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. pp. 189–199. UIST '18, Association for Computing Machinery, New York, NY, USA (Oct 2018). https://doi.org/10.1145/3242587.3242666
16. Patwa, P., Sharma, S., Pykl, S., Guptha, V., Kumari, G., Akhtar, M.S., Ekbal, A., Das, A., Chakraborty, T.: Fighting an Infodemic: COVID-19 Fake News Dataset. arXiv:2011.03327 [cs] pp. 21–29 (2021)
17. Ruchansky, N., Seo, S., Liu, Y.: CSI: A Hybrid Deep Model for Fake News Detection. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. pp. 797–806. CIKM '17, Association for Computing Machinery, New York, NY, USA (Nov 2017). https://doi.org/10.1145/3132847.3132877
18. Secundo, G., Riad Shams, S.M., Nucci, F.: Digital technologies and collective intelligence for healthcare ecosystem: Optimizing Internet of Things adoption for pandemic management. Journal of Business Research **131**, 563–572 (Jul 2021). https://doi.org/10.1016/j.jbusres.2021.01.034
19. Shabani, S., Charlesworth, Z., Sokhn, M., Schuldt, H.: SAMS: Human-in-the-loop approach to combat the sharing of digital misinformation. CEUR Workshop Proc. (2021), https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/pt/covidwho-1208118
20. Shabani, S., Sokhn, M.: Hybrid Machine-Crowd Approach for Fake News Detection. In: 2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC). pp. 299–306 (Oct 2018). https://doi.org/10.1109/CIC.2018.00048

21. Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H.: Fake news detection on social media: A data mining perspective. ACM SIGKDD explorations newsletter **19**(1), 22–36 (2017)
22. Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, R.: World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). International Journal of Surgery (London, England) **76**, 71–76 (Apr 2020). https://doi.org/10.1016/j.ijsu.2020.02.034
23. Tschiatschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A., Krause, A.: Fake News Detection in Social Networks via Crowd Signals. In: Companion Proceedings of the The Web Conference 2018. pp. 517–524. WWW '18, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE (Apr 2018). https://doi.org/10.1145/3184558.3188722
24. Usher, N.: Breaking news production processes in US metropolitan newspapers: Immediacy and journalistic authority. Journalism **19**(1), 21–36 (Jan 2018). https://doi.org/10.1177/1464884916689151, publisher: SAGE Publications
25. Yang, C., Zhou, X., Zafarani, R.: CHECKED: Chinese COVID-19 fake news dataset. Social Network Analysis and Mining **11**(1), 58 (Jun 2021). https://doi.org/10.1007/s13278-021-00766-8