

# Rumor Propagation and Detection Techniques in Social Networks

Zichang Qiu \*

School of Artificial Intelligence, Wuchang University of Technology, Hubei, 430000, China

**Abstract.** Social networks have become the core platform for information exchange in the new era, and their uniqueness is reflected in the extensive interconnections brought about by complex ties and technological progress. These characteristics jointly shape social networks and have a profound impact on the development of the new era. With the popularity of the Internet, the problem of online rumors has become increasingly prominent, bringing many negative effects on society. Due to the fast speed of information dissemination, extensive user participation, difficulty in verification, and significant group effects, online rumors can quickly spread in social networks and have an impact. Therefore, strengthening the supervision of network information is crucial to maintaining justice and order in cyberspace. The four leading rumor detection techniques are currently covered in this article, along with an introduction to the fundamentals of rumor detection. Prospects seem to offer fresh approaches to rumor detection in light of the potential difficulties facing societal development. Analyze the experimental data of different detection methods and explore the advantages and disadvantages of each technique. It is hoped that through in-depth research, information dissemination can be guaranteed, contributing to creating a safe and harmonious cyberspace.

## 1 Introduction

Rumors refer to news without factual basis or hearsay that has not been officially confirmed. Rumors have existed since ancient times and have become particularly prominent in the information age. In the Internet era, rumors vary by their quick spread, broad influence, variety of transmission methods, unclear content, high degree of negative destructiveness, difficulty in countering them, and clear group polarization [1].

Weibo and Twitter are two examples of social networks that have grown significantly in importance as a platform for information sharing and publication due to their ongoing development. Network rumors have proliferated due to social media platforms' ineffective oversight of network information. Network rumors are detrimental to society in many ways. Its impact is not limited to the virtual network space, but also deeply penetrates all aspects of real culture. It is harmful to individuals, culture, economy, industry, social order, social stability, public cognition, values, law, and public security. It not only profoundly affects the

---

\* Corresponding author: [20150775101@mail.sdufe.edu.cn](mailto:20150775101@mail.sdufe.edu.cn)

stability and development of society, but also threatens everyone's daily life and sense of security.

Scholars have done a great deal of investigation and study in response to the current issue of social network rumors. The current state of online rumor detection techniques can be divided into three types: machine learning-based strategies, deep learning-based techniques, and manual techniques, which assess rumor information directly using human intuition and experience. However, they are time-consuming, inefficient, and have high costs. They are also lagging and cannot meet the needs of modern social network rumor detection. Machine learning-based detection methods automatically identify rumors by building a classification model. They are more efficient than manual detection methods, but they still require manual selection and feature extraction, which takes a lot of time and has limited robustness. The resulting feature vectors may not be robust enough. Deep learning models are employed in deep learning-based detection techniques that automatically gather high-level concepts. They are more representative and have lower costs, higher efficiency, and higher accuracy than other detection methods.

In response to the dynamic research and hot frontiers of social network rumors, the academic community has summarized and concluded social network rumor detection [2]. Sudden public events often become the fuse of online rumors. The emergence of new media such as short videos has further broadened the channels for rumor information dissemination, making online rumor dissemination scenarios more diversified, the dissemination content more complex, and involving more fields. After continuous research and exploration, many new theories have emerged, opening up new paths for research in this field. With the widespread application of big data technology, this new technology enables the psychological factors of the masses to be incorporated into the research of online rumor detection during the research process, thus providing a new perspective for in-depth analysis of the generation mechanism and propagation characteristics of online rumors. It will not only help control and govern of online rumors in the future but also provide a valuable reference for subsequent development.

## **2 Internet rumor detection process**

The process of identifying hoaxes on the internet consists of four steps: processing data, selecting and extracting features, training a model, and detecting rumors. The two parts of data processing are data collection and data labeling. Data can be gathered using the software's API, spiders, and publicly accessible datasets. It takes a lot of time and labor to figure out the groups in the first data collection using the conventional labeling technique, which includes paying experts. Wu et al. noted that there are several obstacles to overcome when using traditional supervised learning techniques for the swift identification of rumors, particularly given how labor- and time-intensive the labeling process is [3]. To this end, they proposed an innovative automatic annotation strategy based on semi-supervised learning and designed a new framework called Crosstopic Emerging Rumor detection (CERT). By integrating clustered data, selected features, and training efficient classifiers, this framework effectively overcomes the time delay problem in traditional annotation methods and significantly improves the efficiency and accuracy of rumor detection.

Feature selection is the selection of representative data features from the original data. In machine learning-based detection techniques, feature selection is essential; feature extraction is the process of identifying and extracting features from vast amounts of random information without requiring human labor. Operationally, feature extraction techniques primarily consist of obtaining a single piece of data to extract features; event-level feature extraction extracts potential connections between data by mining hierarchical relationships between data.

Convolutional neural networks (CNN), recurrent neural networks (RNN), graph neural networks (GNN), Transformers, and other deep neural network models are used in detection techniques based on deep learning for social network rumor detection. Propagation trees, aggressive learning, cross-domain techniques, multi-task learning, supervised and semi-supervised techniques, knowledge graph-based techniques, and other techniques are also included. To select an appropriate model for model training based on different scenarios, real-time parameter adjustment is required to achieve the best effect of the model. Rumor detection requires real-time detection, using well-tuned models to monitor data, detect rumors promptly, and ensure the authenticity of the information.

### **3 Online rumor detection methods**

#### **3.1 Manual detection methods**

The manual detection method relies on professionals to directly judge the authenticity of information. This method usually relies on people's professional knowledge and experience to judge the authenticity of information. The advantage of this method is that the accuracy rate is relatively high. Its drawback is that processing big amounts of data has evident efficiency and latency issues, making it challenging to meet the demands of continual tracking of enormous amounts of data.

Sina Weibo has implemented two rumor information detection methods, namely Weibo rumor refuting [4] and reporting processing hall [5]. Weibo rumor refuting is a public account on the Weibo platform. It disseminates misleading information to Weibo users through a public account; the reporting processing hall promptly identifies undetected rumor information through reports from Weibo users and publishes the detection results. Both methods of Sina Weibo use people to detect rumors. The detection accuracy rate is very high, but because it is manually detected, there is a significant lag, and it cannot guarantee that the rumor information will be detected the first time, which cannot meet the needs of the modern network society to detect rumor information in real-time.

#### **3.2 Detection methods based on machine learning**

The machine learning-based approach uses an algorithm model to classify the data and automatically learns its characteristics, treating the myth tracking issue as a matter of binary classification in supervised learning. This technique enhances the automation and effectiveness of detection by automatically learning the characteristics of the data using the framework of algorithms and classifying them. The drawbacks are that it requires a lot of labor, materials, and time, and it depends mostly on human feature extraction and selection. Additionally, the stability of the vectors of features it uses may not be sufficient.

Castillo et al. concentrated on automated techniques for assessing a particular set of tweets' trustworthiness [6]. They analyzed posts related to hot topics and classified them into credible and uncredible based on the features of user posting and forwarding behaviors extracted from them. The evaluation results showed that there were observable changes in the way messages were transmitted, and they could be automatically identified as credible and uncredible, with precision and recall rates between 70% and 80%. Liang et al. Found [7] that picking out machine learning techniques had less of an effect on legend identification accuracy than feature design and selection. To forecast if a post is a rumor, they combined five novel functions based on user usage with already-existing, reliable, and efficient criteria based on user behavior. This allowed them to study rumor identification systems. The findings demonstrated that buzz analysis through mass behavior outperforms detection based

on Weibo's built-in features. To automatically recognize rumors from a collection of genuine and false information, Yang et al. gathered verified false rumors on Weibo [8], looked at a set of comprehensive features that may be derived from Weibo, and educated a classifier. Experiments demonstrated the effectiveness of the new features that were suggested for classification. To reliably identify tumors in real-time, machine learning-based rumor detection relies on the choosing and extraction of significant features to describe data. However, manual feature selection and extraction are labor-intensive and material resource-intensive, and the robustness of characteristic vectors is low.

### **3.3 Detection methods based on deep learning**

Deep network topologies are mostly used by deep learning-based detection techniques to gradually acquire more intricate and complex feature representations to enhance classification performance. The advantage of this method is that it can better mine the intrinsic representation of data compared to traditional machine learning methods. It can automatically learn more complex and abstract feature representations through deep network structures, effectively improving classification performance. Its disadvantage is that in practical applications, how to choose a suitable network structure and avoid parameter adjustment are still challenges.

Zheng and colleagues have developed a Transformer-based social media rumor identification model that does not take into account text position data. They put up a better Transformer model that is capable of extracting text location data efficiently [9]. While absolute motion encoding is used to relate words in different places to distinct feature spaces, an achievable contextual position encoder is used to capture the text's direction and distance information. The experimental results demonstrate that the improved model's accuracy on the Twitter15, Twitter16, and Weibo sets of information has risen by 0.9%, which is 0.6%, and 1.4%, respectively, when compared to other baseline models. Based on the rumor transmission path, Zhang et al. collected text as user credibility features and built a multi-feature rumor recognition model based on the transmission tree [10]. This model aggregates text propagation information through a graph convolutional network, uses a multi-head attention mechanism to mine the intra-layer dependencies of the text propagation tree, and constructs a credibility sequence for each user in the user propagation tree. The M-Attention module is used to capture effective user credibility features. Based on experimental results, the suggested method outperforms the present efficient dissemination trees model Binary Graph Convolutional Network by 4.8%, 4.2%, and 3%, respectively, with detection accuracy of 89.3 percent 91.7 percent, and 96.4 percent on Twitter15, Twitter16, and Weibo datasets. A deep focus model constructed using the use of recurrent neural networks was suggested by Chen et al. [11], which learns a temporal neural representation of successive posts selectively to identify rumor information. A multitude of experimental results demonstrate that the proposed model is more accurate as well as faster than other equivalent models. It incorporates soft consideration into the repetition and lumps various characteristics with a specific focus on producing hidden arguments, which may represent the surrounding changes of the post-time. Although the deep learning rumor detection model is far more accurate and efficient, it is difficult to comprehend and requires less labor and material resources.

## **4 Future Outlook**

Research on trend spotting in online communities is still very important. The complexity and significance of rumor detection are increasing with the ongoing development of self-media. So far, the academic community has developed three major types of rumor detection methods, and there is still a need to move forward in the future. In terms of technology, we need to use

the enhancement of computing resources and the optimization of algorithms to make deep learning models more complex and accurate, in order to process bigger data sets and raise the rumor detection accuracy; in the future, more research will focus on how to integrate data from different modes to make detection more comprehensive; to increase the transparency and dependability of the detection findings, the model's interpretability must be improved; the application scenarios cannot be limited to traditional media platforms, but should be expanded to emerging self-media platforms to make the coverage wider; real-time detection of rumors, monitoring the spread of rumors and stopping them in time at the first time, helping the public identify and avoid rumors; using big data to analyze user browsing preferences and conduct personalized rumor filtering, reduce the impact of rumors on individuals; in terms of social impact and ethical considerations, when doing rumor detection, it is imperative to preserve user privacy and refrain from violating it; strengthen public education in society, improve the public's media literacy and critical thinking ability, and help the public better deal with rumors; the characteristics and patterns of rumor propagation may vary in different languages and cultural backgrounds; future research can carry out cross-language and cross-cultural research to adapt to rumor detection tasks under different environments and needs; due to the complexity and extensiveness of rumor propagation, future research needs to build a multi-party cooperation mechanism, including government, enterprises, academia, etc., to jointly deal with the problem of rumor information in modern social networks.

## 5 Conclusions

The pace and reach of rumor dissemination have significantly increased due to social networks' rapid development and extensive implementation. Social network rumor detection still needs to be further studied. Research in the field of rumor detection should develop in a more intelligent, precise, and diversified direction. Against the background of the rapid development of self-media, the prompt and precise identification of rumor information is essential for the present network environment regulation and upkeep. The main focus of this paper is the social networking site rumor detection approach. It presents the platform's rumor detection process, examines and summarizes the benefits and drawbacks of the detection technologies already in use, and discusses the issues associated with rumor monitoring going forward. To sum up, to advance social media rumor detection in the future, researchers must constantly focus on technological innovations while also being mindful of the difficulties and moral dilemmas that arise in real-world uses. Social network rumor tracking will become more important as technology advances and work together to create a clear and safe network environment.

## References

1. Y. Gao, G. Liang, F. Jiang, et al., A review of rumor detection in social networks, *J. Electron.*, **48**(7), pp. 1421, 2020.
2. C. Xu, Research trends and hot topics of online rumors: Visual quantitative analysis based on CiteSpace knowledge graph, *Oper. Res. Fuzzy Theory*, **13**(5), pp. 4926–4939, 2023.
3. L. Wu, J. Li, X. Hu, et al., Gleaning wisdom from the past: Early detection of emerging rumors in social media, in *Proceedings of the 2017 SIAM International Conference on Data Mining*, pp. 99–107, (2017).
4. X. Yang, Analyzing the new model of Weibo rumor-busting by taking the COVID-19 rumor-busting method as an example, *Sound Screen World*, **15**, pp. 116–117, 2020.

5. M. Hou, R. Gao, The generation of information epidemic in social media and the mechanism of rumor refutation: Taking Sina's "Weibo rumor refutation" as an example, *Journalism Commun.*, **22**, pp. 39–41, (2022).
6. C. Castillo, M. Mendoza, B. Poblete, Information credibility on Twitter, in *Proceedings of the 20th International Conference on the World Wide Web*, pp. 675–684, (2011).
7. G. Liang, W. He, C. Xu, et al., Rumor identification in microblogging systems based on users' behavior. *IEEE Trans. Comput. Soc. Syst.*, **2**(3), pp. 99–108, 2015.
8. F. Yang, Y. Liu, X. Yu, et al., Automatic detection of rumor on Sina Weibo, in *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics*, pp. 1–7, (2012).
9. H. Zheng, Y. Hao, H. Yu, et al., Social media rumor detection based on improved Transformer. *J. Netw. Inf. Secur.*, **8**(4), pp. 168–174, 2022.
10. X. Zhang, S. Pan, Q. Mao., Multi-feature rumor detection method based on propagation tree. *J. Electron.*, pp. 1–10, 2024.
11. T. Chen, X. Li, H. Yin, et al., Call attention to rumors: Deep attention-based recurrent neural networks for early rumor detection, in *Proceedings of the Trends in Applied Knowledge Discovery and Data Mining: PAKDD 2018 Workshops*, pp. 40–52, (2018).