Research on Weibo Rumors Recognition: A Survey

Qinying Wang

School of Business School, Shandong Normal University, Jinan 250014, China.

1481623217@qq.com

Abstract

With the development of social media, social media platforms such as Facebook, Twitter, Sina Weibo and WeChat have become important channels for people to share and obtain information. The rapidly spreading rumor information has become an increasingly prominent problem. How to quickly and accurately identify rumors It is an important prerequisite for timely and effective blocking of rumors. This paper summarizes the microblogging rumor recognition work. Firstly, it introduces the relevant academic research on rumors. Then, it introduces the main work and methods of rumor recognition. Finally, it summarizes the development trend of rumor recognition and authenticity judgment.

Keywords

Rumors recognition, characteristics, authenticity.

1. Introduction

With the development of mobile Internet and social media, social media such as social networking sites, blogs, and Weibo have become important sources for users to share information [1]. People are increasingly relying on social media to understand the occurrence of emergencies and the progress of events. In this case, many users will post information on social media, and the reliability of this information can‘t be verified immediately, which leads to the emergence and spread of rumors. For example, the 2011 Fukushima nuclear power plant leakage accident in Japan, there are rumors that Japan's nuclear radiation will pollute seawater, resulting in the production of salt can not be used later, and eating iodine-containing edible salt can prevent nuclear radiation, the news on Weibo and WeChat circle of friends was madly forwarded, which led to a large number of people buying salt.

Sina Weibo and Twitter are the most popular, most influential and most user-friendly social media platforms at home and abroad. Weibo texts generally do not exceed 140 words. The information forms are diversified and can be text, pictures and audio. Users can The content is published in real time. However, due to the limitation of the word size, Weibo content sometimes cannot fully express the facts. The threshold for information dissemination is low, and any registered user can post, comment on and forward other users' Weibo information to express their personal opinions and feelings. However, the authenticity of the information published by ordinary users is not authoritative. Some are just hearsays but exaggerate the facts and publish them on the Internet. The characteristics of fissile propagation speed will make the rumor information spread rapidly and can trigger the society in the first time. Hot issues, quickly gathered the focus of public opinion.

In order to solve the problem of the proliferation of Weibo rumors, researchers and Weibo operators have made various efforts. For example, Sina Weibo uses crowdsourcing technology to establish a microblog community management center, launches the “microblogging hack” official account [2], and conducts 24-hour real-time monitoring of Weibo information published by users in Sina Weibo, according to various official The information collected by the authoritative information channel identifies and judges the content posted by the user, and then publishes the rumbling result to the

official account. Similarly, Hukou.com established a group of rumors and shredders to build rumor shredder accounts on Sina Weibo and regularly publish rumors. These rumors have played a certain role in suppressing the spread of rumors, but they require a lot of manpower, material resources and financial resources. At the same time, rumors are lagging behind, often after the rumors have been flooded, the rumors can be released. However, the bad consequences are here. It may have already appeared. Therefore, many researchers set out to carry out the automatic detection of microblogging rumors, by exploring the key factors of rumor communication, in order to discover the distinctive features or internal laws of rumor communication, and to build models to improve the automatic identification of microblog rumors through improved algorithms. Thus to curb the spread of rumors before the massive outbreak.

2. Research on Rumors

The academic study of rumors began in the period of World War II. Due to the necessity of confidentiality of military operations, people are not able to obtain relevant information about national security in a timely manner. These are closely related to the life and security of the people, and the emergence of rumors just fills up. This part of the gap expressed the people's desires, fears and hostility, so the rumors grew during the war.

In 1944, Knapp published an article "A Psychology of Rumor", which began to systematically describe the field of rumors from the causes and consequences of rumors [3], and focused on the issue of rumor control, according to the emotional needs expressed, Knapp divides rumors into three categories: Wish Rumor, Bogie Rumor, and Aggression Rumor. In 1947, Allport and Postman [4] proposed the rumor formula R=I×A in the book Proverbs Psychology, where I (Importance) represents the importance of the event, and A (Ambiguity) represents the expression of the event. Fuzziness establishes a close relationship between rumors and subjective emotional states. If one of the importance or ambiguity is 0, then the rumors will not spread. Allport and Postman's definition of rumors is "a special or current current-related statement that aims to convince people that it is generally circulated among people in a word-of-mouth manner, but lacks specific information to confirm its certainty." This definition clearly states that information has a certain degree of ambiguity before it becomes a rumor and spreads.

Kapferer believes that the definition of rumors with "unconfirmed" is logically unreasonable [5]. Emphasizing rumors is "unsubstantiated" news, which belongs to the most intense historical stage of suspicion of rumors, so this is only a definition of ideology. Therefore, in the book "Proverbs: The World's Oldest Media", Kapferer defines rumors as information that has appeared in society and has not been officially publicly confirmed or has been rumored by the government. He believes that rumors are both A social phenomenon is also a political phenomenon.

In the 1990s, after the advent of the World Wide Web, researchers began to study rumors on the Internet to understand how rumors were generated and disseminated on the Internet, trying to discover the laws of rumors, so that they could be rumored before the massive rumors broke out. And block. Zanette studied the dynamic propagation model of infectious diseases on the small world network [6]. For the first time, the evolution of dynamic processes was compared in small world networks and dynamic small worlds. Buchegger et al. studied the propagation of rumors in mobile peer-to-peer networks [7], proposed a Bayesian method for reputation characterization, updating and attempting to integrate, and proposed a mechanism to detect and exclude potential rumors. Moreno et al. studied the dynamic propagation of rumors in complex networks [8], and obtained the mean field equations describing the kinetic characteristics of rumors. These equations were numerically solved by random methods.

According to the evolution of the media, the study of rumors is roughly divided into three stages: the first stage is traditional media, the main media are newspapers, radio and television, and the word-of-mouth between people, chained Communication is the main, the transmission speed is slow,

the scope is small, and the information source is relatively single. The research at this stage mainly focuses on the generation, dissemination and influencing factors of rumors, analyzes the psychological motives of the rumors, and lays a huge and solid foundation for the follow-up research. Foundation, and formed a relatively complete system of disciplines, but it is difficult to find real data in natural life to reflect the real situation of communication in reality. The second stage is computer media, which relies mainly on e-mail, portal websites, and instant messaging software. The speed of communication and the scope of influence have been greatly improved. The channels for obtaining information are diversified. Researchers use computers to model network features. Analysis, in order to discover the law of rumors in the network. The third stage is social media, which mainly includes social software and forums such as Weibo, WeChat, QQ, etc. The information dissemination method is mainly web-based communication, the speed of dissemination is fast, the scope of influence is wide, and the interaction between users is fully reflected. At this stage, the rumor features and propagation laws are mainly explored through big data and machine learning algorithms, and the automatic recognition and monitoring of rumor information is realized, and it is smashed in the germination before the rumors are overwhelming.

3. Research on Weibo Rumors Recognition

The existing rumor recognition mainly adopts the classification method, and regards rumor recognition as a two-category problem. By extracting relevant features, such as user features, communication features, text features, etc., natural language processing, social network analysis, data mining, and machines are used. Learning techniques and method selection can clearly distinguish the feature set of rumors and non-rumors, and then use different classification algorithms to construct classifiers, such as decision trees, Bayesian, support vector machines and random forest classification algorithms, according to known training. The set uses the selected classification algorithm for learning, and the training results in a classification model. There are only two classification results: the information is a rumor, and the information is not a rumor. The evaluation of the model generally uses accuracy, accuracy, recall and F1 values.

A. H. Wang implemented a spam detection prototype system to identify suspicious users on Twitter [9], and proposed a directed social graph model to explore the relationship between "followers" and "friends" among Twitter users, based on User relationship graph and text-based content are collected from the data published by Twitter, and then Bayesian classification algorithm is applied to distinguish between suspicious behavior and normal behavior. The experimental results show that the accuracy of the system detection is 89%. Ratkiewicz developed a "Truthy" system based on Twitter [10], introducing an extensible framework for real-time analysis of event diffusion in Twitter by mining, visualizing, classifying and modeling public microblog events, tracking Twitter. Political rumors, although the accuracy of information classification is very high, but this system is more inclined to emotional analysis.

Based on the Ratkiewicz study, V. Qazvinian first classified the collected Tweets [11], and verified the validity of the three characteristics of text content, network features and microblog features in rumor recognition. A Bayesian classifier and an integrated classifier were built on the dataset to simulate more than 10,000 manually annotated tweets collected from Twitter to analyze the user's trust behavior with these rumor-related Tweets. It shows that the average accuracy MAP of the search model has reached more than 0.95, which opens up a new dimension for analyzing online rumor information and other aspects such as microblogging sessions.

Castillo extracted 68 features from Twitter [12] and classified them into the following four categories: content-based features, user-based features, propagation-based features, and topic-based features, using the best selection method to select 15 of the most Good features are used for the construction of classifications, including the average emotional score of the message, the proportion of messages with positive emotional scores, the proportion of messages with negative emotional scores, the proportion

of messages with URLs, the number of different short URLs, and the most frequently posted authors. The proportion of the message, the proportion of messages containing the user @ mentioned, the proportion of messages containing the question mark, the proportion of messages containing smiley symbols, the proportion of messages containing the first person, the average age of registration of the user, the average number of fans of the user, the average number of friends of the user, The average number of messages published by the user and the maximum number of nodes included in the layer other than the child of the root node, and nearly one-third of the features are related to the emotions expressed by the user. They used the J48 decision tree classification algorithm to build the classifier, and the accuracy rate was 86% when predicting whether the news topic was credible.

Takahashi conducted an example analysis of the Twitter rumors caused by the Japanese tsunami [13], studied how rumors spread after the earthquake disaster, and discussed how to deal with the rumors. After the disaster, the actual situation of the rumors was first investigated, and then the characteristics of these rumors were found. It was found that the explosion point, forwarding rate and word distribution difference were helpful for detecting rumors, and based on these characteristics, a rumor detection system was designed, which can be detected from Twitter. Information that may be rumors is evaluated and judged.

Zubiaga et al. believe that marking unauthenticated messages may not avoid the dissemination of information [14]. The detection of rumors can also form a rumor tracking system that finally determines its authenticity, and proposes a new rumor detection method. Learn from the social media's sudden news dynamics to discover rumors of new events, a total of 5,802 Twitter messages, 34% of which are rumors and 66% of which are non-rumors. They collected the Twitter data of five major event news reports, using the conditional random field as the sequence classifier. Instead of querying each piece of information, it uses the learning context in the event to detect whether it is a rumor, thus enhancing the classifier. Suitability, where the best timeline F1 score is increased by nearly 40%. In addition, Zubiaga et al. proposed a method [15] to collect, identify and annotate the datasets of a total of 4,482 tweets of 330 rumors related to 9 news events, and analyze the dataset to distinguish the two states of the rumor life cycle. , before and after clarifying its authenticity, to understand how users spread, support or deny rumors that were later confirmed to be true or false, and finally come to a conclusion that true rumors are often resolved faster than wrong rumors, and users The general tendency is to support every untested rumor.

Yang et al. used the rumor data [16] obtained from Sina Weibo to extract two new features from the message, one is the client used by the user to publish the message, and the other is the location of the event mentioned in the message content. They found that about 71.8% of the rumors were posted on non-mobile clients, and about 56.1% of the events mentioned in the rumor message content occurred abroad. In order to verify the validity of these two features for identifying rumors, combined with the existing content-based, user- and communication-based features, the SVM classifier is constructed for rumor recognition. Zhang et al. believe that shallow features do not distinguish between rumors and non-rumors in many cases [17]. In order to solve this problem, they extracted four implicit features from the content, including popularity orientation, internal and external. Consistency, emotional polarity and commentary perspective, based on selected features to build SVM classifier, the accuracy of recognition reached 72.4%

For the first time, Jin et al. attempted to identify rumors by establishing a propaganda model of rumors on Twitter [18], using the enhanced infectious disease model SEIZ (Susceptible- Exposed- Infected- Skeptic) for eight types of rumors from different parts of the world. Modeling is performed to obtain conversion parameters between user states according to the model, and the ratio forms of the parameters are used to identify rumors and news. Although the degree of discrimination for some special topics is not very high, opening up new ideas for rumor recognition in combination with the law of transmission has yet to be further studied.

4. Conclusion

The rumor recognition and detection work in Weibo involves many aspects such as information content, network characteristics and human behavior and psychological characteristics. This is not only a text recognition work or a network research work, but also a human Social psychology research. Therefore, the rumor recognition work should comprehensively consider the impact of these characteristics, build a model, system or algorithm based on full consideration of various factors, in order to discover the law of rumor propagation, real-time detection of rumors, and thus as much as possible Reduce the harm caused by the proliferation of rumors. So far, the identification of rumors involves theories, techniques and methods in many fields such as natural language processing, social network analysis, sociology, psychology, machine learning, data mining, etc., so this is an interdisciplinary task that strengthens various disciplines. The integration and cooperation between them is of great significance and value.

At present, feature mining and law discovery work from microblog content, users and communication, modeling based on the discovered feature rules, the accuracy of rumor recognition has reached more than 80%, but still not found in rumors Significant characteristics and basic laws in the process of communication, so we can consider further research on the law in the process of rumor information dissemination. At the same time, the purpose of real-time detection of Weibo rumors is to be able to detect in the early stage of rumors, to curb its explosive trend and reduce the harm and adverse effects brought about by rumors. Therefore, the real-time identification and detection of rumors will be the focus of future research and application.

References

[1] M. Mendoza, B. Poblete, C. Castillo. Twitter under crisis: Can we trust what we RT. Proceedings of the First Workshop on Social Media Analytics, (2010), 71-79.
[2] http://weibo.com/weibopiyao?c=spr_qdhz_bd_baidusmt_weibo_s&nick=&is_hot=1
[3] Knapp R H. A psychology of rumor. Public Opinion Quarterly, vol. 8 (1944) No. 1, 22-37.
[4] Allport G. W. and Postman L, An Analysis of Rumor, Public Opinion Quarterly, Vol. 10 (1946-1947), 3-36.
[5] Kapferer J N. Rumors: World's Oldest Media. Paris, France: Le Seuil Editions, (1987).
[6] Zanette D H. Dynamics of rumor propagation on small-world networks. Physical Review E, Vol. 65 (2002) No. 4, 041908.
[7] Buchegger S, Le Boudec J. The effect of rumor spreading in reputation systems for mobile ad-hoc networks. Proceedings of the 8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks. Sophia-Antipolis, France, (2003), 1-10.
[8] Moreno Y, Nekovee M, Pacheco A F. Dynamics of rumor spreading in complex networks. Physical Review E, Vol. 69 (2004) No. 6, 066130.
[9] A. H. Wang, Don't follow me: Spam detection in Twitter, Athens, Greece: IEEE, (2010).
[10] Ratkiewicz J. Conover M, Meiss M, et al. Detecting and Tracking the Spread of Astroturf Memes in microblogs, Edinburgh, United Kingdom: Association for Computational Linguistics, (2011).
[11] Qazvinian V, Rosengren E, Radev D R, et al. Rumor has it: identifying misinformation in microblogs, Edinburgh, United Kingdom: Association for Computational Linguistics, (2011).
[12] Castillo C, Mendoza M, Poblete B. Information credibility on twitter, Hyderabad, India: ACM, (2011).
[13] Takahashi T, Igata N. Rumor detection on twitter. Kobb: IEEE, (2012), 452-457.
[14] Arkaitz Zubiaga, Maria Liakata, Rob Procter. Learning reporting dynamics during breaking news for rumor detection in social media, arXiv preprint arXiv: 1610. 07363, (2016).
[15] Arkaitz Zubiaga, Maria Liakata, Rob Procter, et al. Analysing how people orient to and spread rumors in social media by looking at conversational threads. PLoS ONE, Vol. 11 (2016) No. 3, 1-29.

[16] Yang F, Liu Y, Yu X, et al. Automatic detection of rumor on Sina Weibo. Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics, Beijing, China, (2012), 13-20.
[17] Zhang Q, Zhang S, Dong J, et al. Automatic Detection of Rumor on Social Network. Proceedings of the 4th CCF Conference on Natural Language Processing and Chinese Computing. Nanchang, China, (2015), 113-122.
[18] Jin F, Dougherty E, Saraf P, et al. Epidemiological modeling of news and rumors on twitter. Proceedings of the 7th Workshop on Social Network Mining and Analysis. Chicago, USA, (2013), 8:1-8:9.