Prediction and Analysis of Rumour's Impact on Social Media

Jun Yin†, Shaowu Liu†, Qian Li†, Guandong Xu†*
† Advanced Analytics Institute, School of Computer Science
University of Technology Sydney.
Email: Jun.Yin-2@student.uts.edu.au, {Shaowu.Liu, Qian.Li, Guandong.Xu}@uts.edu.au

***Abstract*—Rumour, as an important form of social communication, has been run through the whole evolutionary history of mankind. People maliciously disseminate rumours in order to increase awareness, slander others or cause panic, *etc*. To eliminate this issue, many researchers resort to detecting rumours on social media. However, rumour detection is not sufficient to eliminate the negative impact, which also requires official institutions to provide the refutations. In practice, the number of rumours on social media is too large, there is no need to refute some rumours with little or no concern. Therefore, we need to evaluate the impact of the rumours in advance. In this paper, we devise a rumour influence prediction model *RISM* (Rumour Impact on Social Media) based on a popular rumour intensity formula to predict the impact of a newborn rumour. Extensive numerical experiments are carried out on the real rumour data that are collected from *Toutiao.com*, which demonstrate the effectiveness of our proposed *RISM* model.**

***Index Terms*—Rumour Impact, Prediction Model, Rumour Analysis, Social Media**

I. Introduction

The last decade has witnessed the rise of social media along with the rumours circulated on it. The wide spread of rumours can not only damage the genuine information but also mislead the public opinions. Consequently, the detection and measurement of rumours become vital for maintaining a healthy social media environment.

Nevertheless, rumours are not a new concept, but has existed for a long time since the printing press was invented in 1439. Despite the long history of rumours concept, there is yet no agreement on the definition of the term "Rumours". Recent publications in the research literature present two factions about the definition of rumours. For the first faction, some recent work misdefined rumours as an item of information that is deemed false [1], [2], mixing up with fake news. While for the second faction, which is the majority of the literature, they defined rumours instead as "unverified and instrumentally relevant information statements in circulation" [3]–[5]. In this paper, we adopt the definition of rumours in line with the second faction, which is also consistent with the definition given by major dictionaries. The Oxford English Dictionary defines a rumour as "a currently circulating story or report of uncertain or doubtful truth"[1]; the Merriam Webster Dictionary defines it as "a statement or report current without known authority for its truth"[2].

Definition 1. *Rumour is an item of information that are unverified at the time of posting, and may turn out to be true, or partly or entirely false; alternatively, it may also remain unresolved.*

Based on the unverified character, it is crucial to verify the authenticity of rumours. However, the increasing use of social media platforms for information and news gathering, its immoderate nature leads to the emergency of a large number of rumours. Therefore, it is not feasible to verify the authenticity of each rumour on social media. Besides, from the perspective of refutation, pay much more attention on the rumours that have little or even no concern is not cost-effective.

To alleviate this issue, we need to filter the rumours before taking the next step. For rumours with a higher impact on social media, we need to pay much more attention to check their authenticity, and post the refutations accordingly. On the contrary, we might not have to specifically check its authenticity. Thus, another important research problem pops up. How should we filter the rumours when they arise? Rumours with a wide spread of trends usually have their own narrative style to attract publics' attention, such as scientific narrative style, star effect style and so on. For example, in some rumours, both the data and the pictures are based on facts, and each data is marked with reference materials or come up with sentence like "According to relevant research...". And this is the means of using data to blind the eyes of the public. If you delve into it, you will find that these so-called relevant research is groundless. Then, do we need to spend time to delve into it? Another kind of rumours uses popular keywords to attract the public's attention and achieve the purpose of spreading on social media, like the star effect. For instance, a news posted in 2017 on *Toutiao.com* (i.e., a news platform in China) said that Lu Han and Guan Xiaotong, two popular Chinese super stars with lots of fans, would have a cooperation stage in the CCTV Spring Festival Gala Evening of 2018, and even attached photos of rehearsals. This news was confirmed to be fake but received a large number of sharing and reposts on social medias. This is a typical use of the star effect to spread rumours among the public. Fortunately, this news didn't have

*corresponding author

[1] https://en.oxforddictionaries.com/definition/rumour.

[2] http://www.merriam-webster.com/dictionary/rumor.

any serious impact, but only disappointed some fans. However, if the star effect is used to spread rumours that may cause loss of public property or even danger to life, it will have a very serious impact on social society. The proposed *RISM* model will give these kinds of rumours a higher impact scores to raise people's attention that they need a further confirmation.

Most of researchers have engaged in the detection of rumours. Based on their work, we aim to address two further challenges as below.

1) How to define the impact of rumours on social media?
2) How to predict the possible impact of rumour at its early stage?

Accordingly, we propose a novel prediction model *RISM* that learns rumour impact on social media. Our main contributions are summarized as follows.

- While related literature is limited, we provide a novel measurement on the impact of rumours.
- A content-based model *RISM* is proposed, which can detect the impactful rumours before being spread.
- We conduct extensive experiments on real-world datasets to demonstrate the effectiveness of *RISM* model.

The rest of this paper is organized as follows. Related work is discussed in Section II; Our *RISM* model is proposed in Section III, which consists of rumour impact measurement and content-based feature extraction; Section IV uses real-word dataset from *Toutiao.com*, a news content platform in China, to demonstrate the validity of our proposed model; and at last, a conclusion and some future work are given in Section V.

II. Related Work

Rumour on social media has existed since news started to propagate through social media. Almost at the same time, researchers took advantages of the opportunities that social media provides to analyse how users discuss or even share rumours [6], [7], and finally detect rumours on social media.

Most of the rumour detection researches are to train a classifier by a set of prelabeled rumours, like Hamidian and Diab [8], [9]. For example, if the prelabeled rumours contain content like "Germinated potatoes are not poisonous and can be eaten", then any news related to it (e.g., "It is safe for sprouted potatoes to be cut off the burgeon and cooked.") will be classified as rumours. Zubiaga $et\ al.$ categorised this kind of rumours as long-standing rumours. Long-standing rumours usually circulate on social media for a long periods of time and have some related rumours known as priori [10].

However, when it comes to newly emergency rumours with no prior information, solely relying on the content similarity between new rumours and priori would not meet the needs. To alleviate this issue, Zhao $et\ al.$ [11]made an assumption that rumours will provoke tweets from skeptic users who question or enquire about their veracity. In another word, if a piece of information has many related enquiry tweets, it means that the tweet is a rumour. Zhao $et\ al.$ then created a manually curated list of five regular expressions to identify enquiry tweets. These enquiry tweets are then clustered by similarity, and tweets in each cluster being ultimately deemed candidate rumours. Furthermore, Zubiaga $et\ al.$ [12], [13] proposed another method to learn the background of entire breaking news story to estimate whether the tweet will become a rumour or not. The methods that proposed by Zubiaga $et al.$ are based on such a hypothesis that without the fully understand of background, we may not be sufficient to know whether the potential story under the tweets is a rumour or not. From another point of view, McCreadie $et al.$ studied the feasibility of using crowdsourcing platforms to identify rumours and non-rumours on social media. This identification of rumours obtains a high-level of consensus among annotators [14].

Although the research on rumour detection is in full swing, there is quite a few research on the impact of rumours. Some social media providers even hire senior journalists working 24h and 7d every week to maintain an official account [15], which exposes new rumours regularly, in order to minimise the negative impact of rumours on their platform (e.g., @*WeiboPiyao* in Sina Weibo). To the best of our knowledge, most existing works regarding rumour impact are solely based on prior knowledge or various other assumptions or even human power. Hence, being targeted to numerically describe the rumour impact on social media and thus help government to control social rumours are now a top priority.

III. Methodology

This section proposes a *RISM* model that can predict the impact of a new rumour, which thus provides a basis for rumours filtering.

A. Problem Statement

Let D be a rumour dataset, consisting of N rumour news $\{d_i\}_{i=1}^{N}$, while each news $d_i = \left\{w_1^i, ..., w_{P_i}^i\right\}$ contains P_i words. Let $H_i = \{h_j\}_{j=1}^{k}$ be a set of k comments related to the rumour news d_i, where each comment $h_j = \left\{w_1^j, ..., w_{Q_j}^j\right\}$ contains Q_j words. We aim to learn a rank list RI based on all sentences in $\{d_i\}_{i=1}^{N}$. Rumour's impact score represents the degree of negative effects caused by rumours. In other words, if a rumour news d_i is predicted to be a higher impact rumour, then, the government or some official institutions need to take measures to refute this rumour officially.

B. Measuring Rumour Impact

The rumour intensity formula was first proposed by American sociologists G.W. Allport and L. Postman in 1947 [4]:

$$R = I * A \tag{1}$$

where R represents the impact of the rumour, I represents the importance of the information mentioned in the rumour and A represents the ambiguity of the rumour.

Dutch scholar Chorus believes that the intensity of rumour is not only related to events, but also includes human factors. Thus, he introduced the concept of audience judgment ability in 1953 [16]. Chorus believes that audience judgment should include personally relevant knowledge, observation and moral cultivation, which are negatively correlated with rumour circulation. In other words, the richer the individual's knowledge is, the stronger the observation is, and the higher the moral cultivation is, the more resistant the spread of rumours is. Therefore, he developed the rumour intensity formula as below.

$$R = I * A/C \tag{2}$$

where C reflects the public's attitude towards the rumour.

The rumour intensity formula has been further developed by some researchers recently. Through the analysis of public emergencies, Wang [17] proposed his rumour intensity equation as below.

$$R = I * A * J * E \tag{3a}$$

$$E = c * s * \frac{1}{o}(s > 1, 0 < o < 1, c > 1) \tag{3b}$$

where I and A have the same representations as in Eq. (1) and Eq. (2), J represents the public critical ability and E refers to the environmental index. The new variable E includes the communication environment index c (communication) and the political environment index. Political environment index is composed of the political stimulus index s (stimulate) and political transparency o (open-politics). From the practical application point of view, the "Political Environment Index" and the "Communication Environment Index" have no specific measurement standards to give them corresponding values, which weakens the operability of the formula to some extent. Furthermore, Hou [18] improved the rumour intensity of Eq. (4) based on the dataset from Weibo (i.e., a popular social media in China). Hou claimed that rumour intensity has some relationship with the identity of the publisher.

$$R = I * A * (V + f) * \frac{1}{c + w} \tag{4}$$

where V denotes the identity of users, f refers to the number of fans, c represents the publics' critical ability and w represents the publics' willingness. However, Hou didn't propose standard measurement for the identity V and public willingness w, which makes Eq. (4) impractical.

Some researchers [19]–[21] focused on the evolution of the rumour intensity equation in Eq. (2) and took advantage of the investigation of rumours' spread within social media. However, we noticed that existing rumour impact models are derived from specific scenarios, which limits their applications.

In this paper, based on the rumour intensity formula proposed by Chorus (Eq.2), we define rumour impact that is adaptable on most social media platforms. The following part gives three significant definitions, i.e., importance, ambiguity and public critical ability.

1) Importance: If one thing (or a person) may cause rumours, this thing (or this person) is of some importance (so-called "focus events" or "top people"), the "focus events" or the "top people" can be a gimmick of rumour mongers that provokes public attention. Namely, if a rumour attracts amount of concerns (i.e. *thumbs up*, *sharing*), a lot of people are willing to spend their time on this topic or even spread it on social media. In this case, this rumour is of higher importance. The actions *thumbs up*, *sharing* and *thumbs down* show that the rumour has aroused public attention, even though the public do not like it.

Thus, we use the following objective functions to define the importance of a rumour:

$$I_i = Z_{score}(\sum CN_i) \tag{5}$$

where I_i means the importance for rumour d_i, CN_i is the sum of *concern*, *thumbs up*, *thumbs down* and *sharing*, Z_{score} is used to normalise values in order to avoid large value spans.

2) Ambiguity: As we mentioned in Def.1 that rumour is an item of information that are unverified at the time of posting, and may turn out to be true or partly or entirely false or remain unresolved. Some rumours may suffer from this uncertainty, the more they have been refuted, the less clear the truth is. Ambiguity is the major factor affecting this uncertainty. The initiator of ambiguity is the intentional obscuration or even distortion of the truth from the issuer of the rumour. For example, some rumour spreaders only describe a part of the facts, but conceal the other part of the news, causing the public to speculate and suspect the original news. However, different people may have different interpretations to the news, which eventually causes the spread of rumours. For the purpose of adaptable on most social media, we give the assumption that the fewer the number of words are, the higher the ambiguity is, and the higher impact the rumour may have, as indicated by the following equation.

$$A_i = \frac{1}{\log_e(P_i)} \tag{6}$$

where A_i is the ambiguity of rumour d_i, while P_i is the number of words in the rumour, For rumour news $d_i (1 \leqslant i \leqslant N)$, we have $P_i > 1$.

3) Public Critical Ability: In the absence of open and transparent information, the public has no way to give a impartial judgment. Furthermore, if this information is relevant to the public's personal interests, and there is no timely feedback. Then, some people may be led by some "reasonable and gimmick" rumours, because of the lack of calm attitude. At that time, the public prefer to believe it before get confirmation. In particular, when the rumour involves issues such as official corruption, once someone maliciously cook up a story, the public is used as a "secondary passer" and a "loudspeaker". Therefore, the critical ability of the public indeed has some influence on rumour impact. Particularly, the calmer the public's attitude and the stronger the critical ability they have, the weaker the impact of the rumour, vice versa.

Meanwhile, on general social media, comments are the straight ways that reflect the public attitudes toward the rumours. For instance, comments like "it must be fake!" or "Only fool will believe it" indicating that users who read this rumour and leave the comments like these have strong critical abilities. On the other hand, comments like "Really?" or "Don't want this to be true." reflecting that users who read

TABLE I
STATISTICS OF CORPUS

Chinese Corpus	Description
Wikipedia 2019	$1,043,224$ well-structured Chinese words from Wikipedia
News Corpus	2.43 million news, collected from 2014 to 2016, covering $6,300$ media
Baidu Baike	1.425 million pre-filtered, high quality questions and answers from Baidu Baike

this rumour and leave the comments like these have weak critical abilities.

Hence, in this paper, we measured the public critical ability according to their attitudes from comments. We utilized *HowNet*[3] to give each rumour's comments a score. Higher score means that readers have a higher critical ability according to this rumour, while the lower score has the opposite meaning.

Similar to *WhatNet* in English world, *HowNet* is a large language knowledge base for vocabulary and concepts in Chinese (including English). *HowNet* adheres to the idea of reductionism, arguing that vocabulary/word meaning can be described in smaller semantic units. This semantic unit is called "Sememe". As the name "Sememe" implies, it is atomic semantics, that is, the most basic and minimum semantic unit that should not be subdivided. In the continuous process of labelling, *HowNet* gradually built a fine set of sememe system (about 2000 sememe). *HowNet* accumulates semantic information for hundreds of thousands of vocabulary/word meanings based on the semantic system. In this paper, we use *HowNet* to analyze the emotional words in the comments, and finally give rumour $d_i(1 \leqslant i \leqslant N)$ a comment emotional score $C_i(1 \leqslant i \leqslant N)$.

For rumour $d_i(1 \leqslant i \leqslant N)$ in D, we can compute the value of *importance* (I_i), *ambiguity* (A_i) and *public critical ability* (C_i). Based on rumour intensity of Eq. (2) proposed by Chorus (Eq.2),impact score R_i of rumour d_i is

$$R_i = \frac{Z_{score}(\sum CN_i)}{\log_e(P_i) * C_i} \tag{7}$$

C. Content-based Feature Extraction

Since we aim to predict the rumour's impact at its the early stage, However, at the time when the rumour appears, usually there is only rumour content without other related attributes. Therefore, in this paper, we only extract features based on the content of the rumours. Specifically, the features we extracted consist of two parts, one is the *TF-IDF* that is widely used in text mining, and another one is *Word to Vector* that represents the semantic information hidden in the text.

1) TF-IDF: TF-IDF is the acronym for term frequency-inverse document frequency, which is a numerical statistic to reflect how important a word is in a document [22]. It is often used as a weighting factor in searches of information retrieval, text mining, and user modelling. *TF-IDF* increases proportionally to the number of times a word appears in the document and is offset by the number of documents in the corpus that contain the word, which helps to adjust for the fact that some words appear more frequently in general. *TF-IDF* is one of the most popular term-weighting schemes today, and it is said that around 83%[4] of text-based recommender systems in digital libraries use *TF-IDF* as features.

Similar to most of previous work, we take advantages of *TF-IDF* as of our features:

$$\mathit{TF}_{a,i} = \frac{n_{a,i}}{\sum_{P_i} n_{P_i,i}} \tag{8a}$$

$$\mathit{IDF}_a = \log \frac{|N|}{1 + |\{i : w_a \in d_i\}|} \tag{8b}$$

$$\mathit{TF\text{-}IDF}_{a,i} = \mathit{TF}_{a,i} * \mathit{IDF}_a \tag{8c}$$

where $n_{a,i}$ is the number of occurances for word w_a in rumour d_i, and the denominator of Eq.(8a) is the number of words in d_i, in Eq.(8b), $|N|$ denotes the total number of rumours in D, $|\{i : w_a \in d_i\}|$ is the number of rumours where the word w_a appears. To prevent the denominator from being zero, we modify the denominator to $1 + |\{i : w_a \in d_i\}|$. And then, we take the product of *TF* and *IDF* as the value of *TF-IDF*.

In order to extract more valuable words as features, we then calculate the information gain of the *TF-IDF* value for each term in each rumour, and filter out the top 1000 keywords, utilising their *TF-IDF* value as *TF-IDF* features.

2) Word to Vector: When calculating *TF-IDF*, each term is independent, and the possible relationship among terms is unknown. Therefore, we take advantages of *Word to Vector* to characterise the possible relationships among terms.

The dataset used in the experiments is from *Toutiao.com*, which will be shown in section IV-A. To learn the vector representation from words, we exploit Chinese corpus [23] that includes *Wikipedia 2019*, *News corpus* and *Baidu Baike*. Particularly, *Wikipedia 2019* contains 1 million well-structured Chinese words. *News corpus* includes 2.5 million news. *Baidu Baike* consists of 1.5 million answers and questions. Then, we use the trained model to calculate the word vector of each rumour's top 1000 keywords. Statistics information of Chinese natural language processing corpus lists in Table I.

[3] http://www.keenage.com

[4] https://en.wikipedia.org/wiki/Tf-idf

TABLE II
COMPARISON OF FOUR CLASSIFIERS WITH DIFFERENT KINDS OF FEATURES

	Linear Regression	Bayesian Ridge	SVM	GBR
Features	R-squared	R-squared	R-squared	R-squared
TF-IDF	0.722	0.631	0.687	0.719
Word to Vector	0.752	0.601	0.602	0.734
RISM	**0.811**	**0.690**	**0.689**	**0.804**

IV. EXPERIMENT

In this session, we carry out experiments to evaluate the effectiveness of the proposed *RISM* model. First, we use the formula Eq. (7) defined in Section III-B to calculate the impact scores as label for each rumour. Then, we use content-based features *TF-IDF* and *Word to Vector* described in Section III-C to predict the impact score.

A. Dataset

We collect a real-world dataset from *Toutiao.com* that is a news content platform in China. The dataset contains both news content and social context information. News content includes the meta attributes of the rumour (e.g., body text), and social context includes the related user social engagements of rumour items (e.g., user comments, number of sharing, number of thumbs up *etc.*) For each news, domain experts provide the ground truth labels of the rumour or non-rumour. We collect the labeled news with rumours as our rumour dataset D. Statistics of dataset D are shown in Table III.

TABLE III
STATISTICS OF DATASET

Dataset	Toutiao.com
#rumours	$10,548$
AVG #words in a rumour	485
AVG #comments	34
AVG #concern	12
AVG #sharing	136
AVG #thumbs up	278
AVG #thumbs down	139

B. Evaluation Metrics

The aim of *RISM* is to predict the rumour impact of the rumour news. To evaluate the effectiveness of the proposed *RISM* model, we adopt a standard metric coefficient of determination, i.e., R-squared.

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{9a}$$

$$\textit{R-squared} = 1 - \frac{\sum_i (y_i - \widehat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \tag{9b}$$

where the dataset D has n values y_1,...,y_n, y_i for *TF-IDF* features or as a vector $y_i = [y_1, ..., y_n]^T$ for *Word to Vector* features. $\bar{y}$ is the mean of the dataset D. Each value y_i associates with a predicted value $\widehat{y}_i$.

C. Performance Comparison

First of all, we calculate the impact scores for each rumour news d_i in D based on the Eq. (7) illustrated in Section III-B, and use the calculated impact scores as the labels.

TF-IDF values *TF-IDF*$_{a,i}$ are then calculated for each word w_a^i in d_i ($1 \leqslant a \leqslant P_i$, $1 \leqslant i \leqslant N$) based on the equations illustrated in Section III-C1. Meanwhile, in order to extract the most valuable words as features, we calculate the information gain IG_a^i of the *TF-IDF*$_{a,i}$. And then we filter out the top 1000 keywords with higher IG and utilize their *TF-IDF*$_{a,i}$ values as *TF-IDF* features. *Word to Vector* features are then computed using *Gensim*[5] on the top 1000 keywords. In case of receiving a calculation limit, we set the size of each vector as 10.

After defining the features and impact score for each rumour news d_i in D, we split train and test datasets with 10 different random seeds for evaluation on Linear Regression (LR) [24], Bayesian Ridge [25], Support Vector Machine (SVM) [26] and Gradient Boosting Regressor (GBR) [27] models. First, we compare *TF-IDF* features with *Word to Vector* features separately. Then, we combine *TF-IDF* features and *Word to Vector* features together to evaluate the effectiveness of the *RISM* model.

Table II shows the performance of comparison methods on different kinds of features. The combination of *TF-IDF* features and *Word to Vector* features, i.e., *RISM*, outperforms the *TF-IDF* feature or *Word to Vector* feature. *RISM* enables the methods to achieve R-squred value with average 0.7.

V. CONCLUSION

The methods proposed to detect rumours on social media are getting increased recently [28], [29], At the same time, the number of rumours on social media is also in a crazy increasing mode. Just detecting rumours does not substantially solve the impact of rumours on the public, some officially refutation are also needed. However, there is no need to refute every rumour on social media as the number is too large.

[5]https://pypi.org/project/gensim/

Therefore, in this work we raised two research questions regards to the rumours on social media. The first one is "How to describe the impact of rumours on social media?". And the second one is "How to predict the likely impact of social media rumours in the early stages?".

To give solutions to these challenges, we proposed the *RISM* model, which is consisted of two components, the calculation of the rumour impact scores and the prediction of it. Experiments on a real-world dataset collected from *Toutiao.com* have demonstrate the validity of our proposed model.

For future work, we intent to further analyse the language style of rumours, which are predicted with higher impact. In addition, how to make the most effective refutation against such rumours is also a very interesting research direction.

REFERENCES

[1] G. Cai, H. Wu, and R. Lv, "Rumors detection in chinese via crowd responses," in *Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 912–917, IEEE Press, 2014.

[2] G. Liang, W. He, C. Xu, L. Chen, and J. Zeng, "Rumor identification in microblogging systems based on users' behavior," *IEEE Transactions on Computational Social Systems*, vol. 2, no. 3, pp. 99–108, 2015.

[3] N. DiFonzo and P. Bordia, "Rumor, gossip and urban legends," *Diogenes*, vol. 54, no. 1, pp. 19–35, 2007.

[4] G. W. Allport and L. Postman, "An analysis of rumor," *Public Opinion Quarterly*, vol. 10, no. 4, pp. 501–517, 1946.

[5] P. Donovan, "How idle is idle talk? one hundred years of rumor research," *Diogenes*, vol. 54, no. 1, pp. 59–82, 2007.

[6] E. Seo, P. Mohapatra, and T. Abdelzaher, "Identifying rumors and their sources in social networks," in *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR III*, vol. 8389, p. 83891I, International Society for Optics and Photonics, 2012.

[7] T. Takahashi and N. Igata, "Rumor detection on twitter," in *The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligence Systems*, pp. 452–457, IEEE, 2012.

[8] S. Hamidian and M. T. Diab, "Rumor detection and classification for twitter data," in *Proceedings of the Fifth International Conference on Social Media Technologies, Communication, and Informatics (SOTICS)*, pp. 71–77, 2015.

[9] S. Hamidian and M. Diab, "Rumor identification and belief investigation on twitter," in *Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pp. 3–8, 2016.

[10] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *ACM Computing Surveys (CSUR)*, vol. 51, no. 2, p. 32, 2018.

[11] Z. Zhao, P. Resnick, and Q. Mei, "Enquiring minds: Early detection of rumors in social media from enquiry posts," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 1395–1405, International World Wide Web Conferences Steering Committee, 2015.

[12] A. Zubiaga, M. Liakata, and R. Procter, "Learning reporting dynamics during breaking news for rumour detection in social media," *arXiv preprint arXiv:1610.07363*, 2016.

[13] A. Zubiaga, M. Liakata, and R. Procter, "Exploiting context for rumour detection in social media," in *International Conference on Social Informatics*, pp. 109–123, Springer, 2017.

[14] R. McCreadie, C. Macdonald, and I. Ounis, "Crowdsourced rumour identification during emergencies," in *Proceedings of the 24th International Conference on World Wide Web*, pp. 965–970, ACM, 2015.

[15] S. Weibo, "@ weibopiyao," 2015.

[16] A. Chorus, "The basic law of rumor.," *The Journal of abnormal and social psychology*, vol. 48, no. 2, p. 313, 1953.

[17] C. Wang, "Construction and resolution of rumor spreading model of public emergencies," *Journal of Modern communication*, no. 6, pp. 45–48, 2010.

[18] F. Y. Qian Wang, "Allport and postman posed rumor propagation formula improvement and its verification: Rumor analysis of sina micro-blog based on the tourist casualties caused by northeast tiger," *Chinese journal of journalism and communication*, vol. 39, no. 11, pp. 47–67.

[19] R. Fallahpour, S. Chakouvari, and H. Askari, "Analytical solutions for rumor spreading dynamical model in a social network," *Nonlinear Engineering*, vol. 4, no. 1, pp. 23–29, 2015.

[20] S. Bajracharya, "Measures of violence: Rumor publics and politics in the kathmandu valley," *Journal of Material Culture*, vol. 20, no. 4, pp. 361–378, 2015.

[21] B. Doerr and M. Fouz, "A time-randomness tradeoff for quasi-random rumour spreading," *Electronic Notes in Discrete Mathematics*, vol. 34, pp. 335–339, 2009.

[22] A. Rajaraman and J. D. Ullman, *Mining of massive datasets*. Cambridge University Press, 2011.

[23] J. Liu, F. Wu, C. Wu, Y. Huang, and X. Xie, "Neural chinese word segmentation with lexicon and unlabeled data via posterior regularization," in *The World Wide Web Conference*, pp. 3013–3019, ACM, 2019.

[24] G. A. Seber and A. J. Lee, *Linear regression analysis*, vol. 329. John Wiley & Sons, 2012.

[25] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[26] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "Liblinear: A library for large linear classification," *Journal of machine learning research*, vol. 9, no. Aug, pp. 1871–1874, 2008.

[27] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[28] Z. Jin, J. Cao, Y. Zhang, and J. Luo, "News verification by exploiting conflicting social viewpoints in microblogs," in *Thirtieth Aaai Conference on Artificial Intelligence*, 2016.

[29] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *Acm Sigkdd Explorations Newsletter*, vol. 19, no. 1, 2017.