Rumor Analysis &Visualization System

Anjan Pal, Alton Y.K. Chua, and Dion H. Goh

***Abstract*—Rumor messages do not spread in isolation on social media but are followed by the cascades of other messages, namely, counter-rumors and uncertainty-expressing messages. To provide a holistic understanding of how messages spread in wake of a rumor outbreak, the aim of this paper is to propose a framework of an integrated Rumor Analysis and Visualization System. The framework identifies four major modules in terms of the core functionalities of the system: 1) Data Crawling; 2) Data Pre-processing; 3) Data Analysis; and 4) Visualization. The functionalities of these four modules help examine the three types of messages—rumor, counter-rumor, and uncertainty-expressing messages—in the wake of a rumor outbreak. Since visualization enables the dynamics of information contagion to be more effectively presented to the user than text-based formats, the integrated system could be used as a training and deployment toolkit to understand the big picture when authorities need to take decisions and actions to deal with rumors.**

***Index Terms*—information contagion, online rumor, social media, visualization.**

I. Introduction

THE term ‘Online Rumor’ can be defined as a form of electronic word-of-mouth (eWOM) that is either false or speculative within the context of a given incident [1], [2]. Online rumors travel fast particularly during emergencies and social crises which are characterized by severe consequences and information uncertainty [3], [4]. They are generally fueled by social media users who give sensational but spurious information the same weight as news from credible sources [5]. Given that social media technologies provide affordances to share messages easily using one-click plugins such as Facebook’s Share button and Twitter’s Retweet button [6], [7], users have the tendency to share unverified assertions to their network of peers without critical evaluation. As this process is set in motion recursively, the ranks of rumormongers swell and rumors become viral.

However, rumor messages do not spread in isolation on social media. They are followed by at least two other types of messages, namely, counter-rumor messages, and uncertainty-expressing messages [8]-[13]. The term ‘counter-rumor’ is used as an antithetical term to a rumor, which encompasses messages that refute or debunk rumors to present the truth [8]-[10]. Uncertainty-expressing messages refer to those that express doubt and further questions about the veracity of rumors [3], [11]. Thus, the three types of messages—rumor, counter-rumor, and uncertainty-expressing messages— co-exist, and represent crowdsourced perceptions about a rumoring phenomenon in the context of the online setting [8], [14], [15].

Most scholarly attention has been trained on three disjointed lines of enquires in the existing rumor literature. One line of investigation uses mathematical theories to simulate the rumor-spreading process in network topology [16], [17]. Another line leans on the technological paradigm to develop algorithms for rumor detection [18], [19] and rumor containment [20]. The thrust is to detect rumors amid non-rumors. A third line of investigation uses the socio-cognitive paradigm to quell rumors through crisis management and corporate communication strategies [21], [22].

However, works that seek to visualize rumors are hitherto far and few. This is a significant research gap in rumor literature because visualization enables the dynamics of information contagion to be more effectively presented to the user than text-based formats [23], [24]. Translating large data sets into a visual interface can help in identifying trends and cascades of messages at the outset of online rumors.

Hence, the aim of this paper is to propose a framework of an integrated Rumor Analysis and Visualization System. The framework comprises four major modules:

1) Data Crawling;
2) Data Pre-processing;
3) Data Analysis; and
4) Visualization.

The system will draw bona fide data from social media platforms, which will be admitted for further processing to visualize messages and their trends in the wake of a rumor outbreak. The inception, spread and eventual demise of rumors can be traced with the help of these four modules of the integrated system. The system can be used as a training and deployment toolkit to understand the big picture when authorities need to take decisions and actions to deal with rumors.

II. Framework of the Rumor Analysis & Visualization System

As shown in Fig. 1, the proposed framework includes four major modules: Data Crawling, Data Pre-processing, Data Analysis, and Visualization. To illustrate how the four modules work, this paper drawn messages from a case in which the fast-food chain Kentucky Fried Chicken (KFC)

Manuscript received November 20, 2018; revised December, 12, 2018. This work was supported by the Ministry of Education Research Grant AcRF Tier 2 (MOE2014-T2-2-020).

Anjan Pal is with the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore 637718 (e-mail: anjan001@ntu.edu.sg).

Alton Y. K. Chua is with the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore 637718 (corresponding author: phone: (+65) 6790 5810 / 6790 6451; e-mail: altonchua@ntu.edu.sg).

Dion Hoe-Lian Goh is with the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore 637718 (e-mail: ashlgoh@ntu.edu.sg).

ISBN: 978-988-14048-5-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

was wrongly accused of selling rats instead of chicken in one of its branches. Specifically, messages were collected from Twitter for the period of June 13–27, 2015 since this time-frame covered the entire life cycle of the rumor.

The detailed functionalities of the four modules are described in the following four subsections respectively.

A. *Data Crawling*

The function of this first module is to collect data from the Web. In particular, the proposed system needs data related to rumoring phenomena, which would be the primary requirement as input to the system. Multi-pronged approaches can be used to collect event-specific messages [2], [11], [20]. These approaches include the use of different application programming interfaces (APIs) and search interfaces to crawl messages from social media platforms such as Facebook and Twitter. The search terms can include various keywords and phrases related to rumoring phenomena.

For the purpose of this paper, messages related to the rumoring phenomenon were collected from Twitter. Adopting multi-pronged approaches, tweets were collected using Twitter's API and search interface. The search terms included various words and hashtags such as "KFC fried rat," "#KFCRAT," and "#KFCFRIEDRAT" to retrieve the event-specific tweets. Tweets along with their meta-data (e.g., screening name, and number of followers) were captured for further processing. This module is followed by the data pre-processing module.

B. *Data Pre-processing*

This module deals with the data pre-processing tasks that include noise elimination, tokenization, as well as Parts-of-Speech (POS) tagging [25], [26]. The data collected from social media platforms tend to contain a lot of noise. Different filtering techniques can be used to eliminate noise from the data. Tokenization can be employed to break each message into words, phrases and other elements called tokens. Hence, data pre-processing tasks need to be employed on crawled data before they can be meaningfully parsed and processed.

The crawled data were filtered out to eliminate noise in terms of non-English, irrelevant or off-topic tweets. Other tasks included removing tweets that were replies to other users and tweets. Tokens were generated from the collected tweets, and thereafter, textual features were extracted in the form of unigrams, bigrams, and trigram. Stanford Part-of-Speech (POS) tagger were used through a customized program to measure the proportions of various POS in the collected tweets [25], [26]. These pre-processing tasks had not only prepared the dataset for further analysis, but also helped to refine the search terms that were further used to collect event-specific tweets related to the rumor.

C. *Data Analysis*

The function of this module is to identify the three types of messages, namely, rumor, counter-rumor, uncertainty-expressing messages in the wake of a rumor outbreak. After the data pre-processing tasks, each message can be considered as the unit of analysis. A potential set of features can be identified to classify the three types of messages. For this purpose, the related literature [12], [18], [19], [27] can be reviewed to extract several content-based features (e.g., whether a message contains multimedia cue; whether a message contains credence; and whether a message contains emotion) and user-based features (e.g., #followers; #followings for Twitter). Thereafter, messages are required to be coded based on the features identified from the

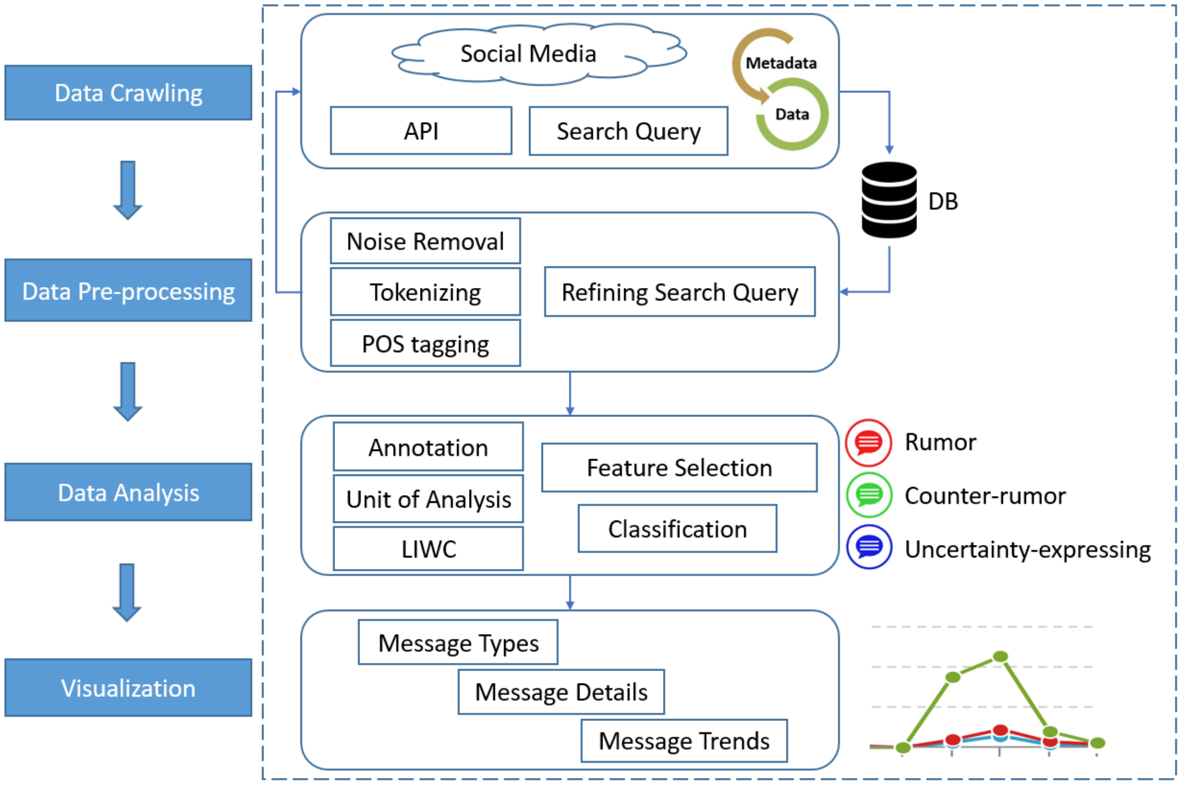

Fig. 1. Framework of Rumor Analysis & Visualization System.
Note. API: Application Programming Interface; POS: Parts-of-Speech; LIWC: Linguistic Inquiry and Word Count.

ISBN: 978-988-14048-5-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

literature. To obtain comprehensive feature set, other linguistic features (e.g., emotional words and tentative words) can also be included from Linguistic Inquiry and Word Count (LIWC) dictionary. Based on the identified feature set, machine learning algorithms such as Naïve Bayes, Random Forest, and Support Vector Machine can be employed to classify messages [12], [18], [19], [27], [28].

For the purpose of this paper, a coding scheme was defined for the three types of messages [3], [9], [29]. As indicated in Table 1, a tweet was identified as rumor when it supported the false claim. In contrast, a tweet was coded as counter-rumor when it refuted the false claim. Moreover, a tweet was coded as uncertainty-expressing message when it expressed doubts and questions about the veracity of the rumor. By employing manual coding, the tweets were coded by two independent coders. Disagreements were resolved through discussion. The average inter-coder reliability in terms of Cohen's k was above 0.70.

TABLE I
MESSAGE TYPES WITH DEFINITIONS

Message types	Coding definitions
Rumor	Tweets supporting the false claim
Counter-rumor	Tweets refuting the false claim
Uncertainty-expressing	Tweets expressing doubts and questions

Fig. 2. Visualizing Message Trends.

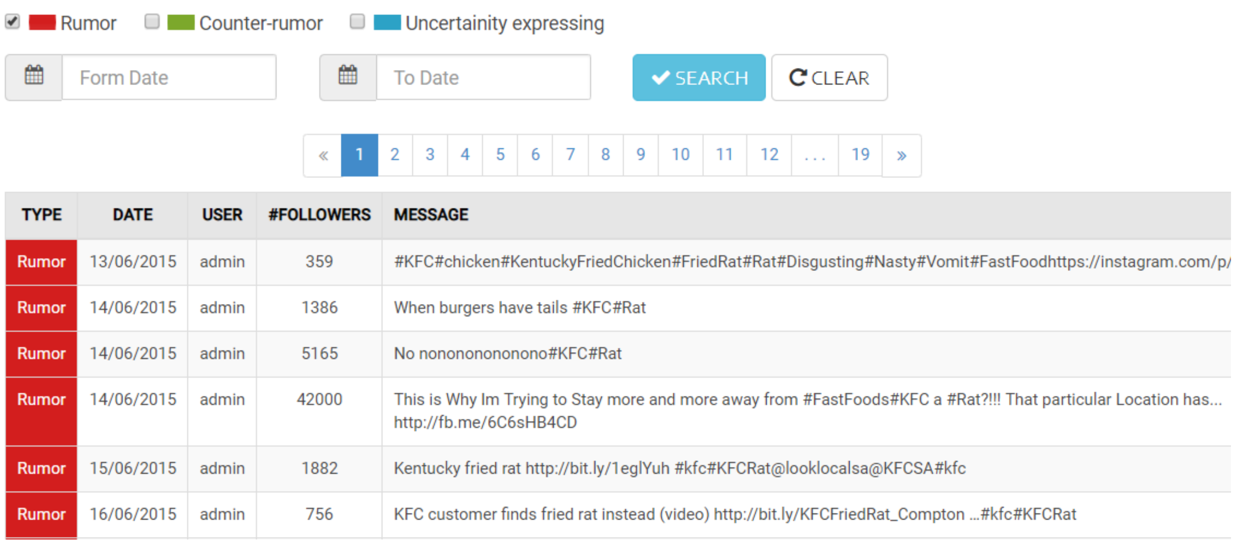

Fig. 3. Message Details Dashboard.
Note. The filed "user" in the message detailed dashboard was displayed with the value "admin" to adhere privacy issue.

ISBN: 978-988-14048-5-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

D. Visualization

This module consists of an interactive interface to visualize messages in terms of their types, details, and trends in the wake of a rumor outbreak. Message types include three different messages: rumor, counter-rumors and uncertainty expressing messages. Message details present the content of the messages with their source and other related information such as date and #retweet. Word clouds are used to visualize the most common terms that appear in three different types of messages. The relative frequency (of occurrence) of each term is indicated by its size within the cloud. Furthermore, the trends of the messages are shown using an interactive play track to visualize the rise and fall of three different types of messages over time.

The interface uses three different colors red, green and blue to indicate the three types of messages, namely, rumor, counter-rumor and uncertainty-expressing messages respectively as shown in Fig. 2. Specifically, the red color connotes the spread of rumor messages as a tell-tale sign of the growing threat. The green color is used to show the sprout of counter-rumor messages. The blue color is used to present the uncertainty involved during the crisis. These polarized color codes help users to interpret the overall picture in terms of visualizing the rumoring phenomenon easily.

To visualize the trends of the three types of messages, an interactive play track was embedded with a graph representing volume of tweets (for each type of message) over time as shown in Fig. 2. It was also synchronized with another visualization panel comprising the three message windows (for the three types of messages respectively) that allows users to visualize the approximate exposure of each tweet. As shown in Fig. 2, each bubble represented a tweet and was sized by the author’s number of followers. For easy reference, each bubble shared the same color as its message type. While a red bubble was used to represent rumor messages, the green and blue bubbles were used to indicate counter-rumor and uncertainty-expressing messages respectively. The interface also allowed users to click on a bubble to view the tweet and its followers.

As shown in Fig. 3, a message details dashboard was used to show the detailed information about tweets. With the selection of a particular message type, the tweets along with other details such as source (users’ screen name) and date would be displayed in the dashboard. The interface further allowed users to select dates in order to facilitate custom selection of tweets.

To delve deeper, Fig. 4 shows the word clouds for the three types of messages. The word clouds helped to depict the most common terms used in the tweets. Moreover, they showed some differences across the three types of messages. For instance, the word cloud for the counter-rumor tweets revealed some of the frequently used terms such as “DNA”, “Test” and “confirm” (related to some events of this chosen case) that were not visible from the word cloud of the rumor tweets. Moreover, the frequency of words and phrases could be obtained from the unigram, bigram, and trigram lists for the three types of messages. An interactive search function also allowed users to tap on a particular word or phrase to check their frequency as shown in Fig. 5.

III. Conclusion

This paper proposes the framework of a Rumor Analysis and Visualization System, which comprises the four major modules, namely, Data Crawling, Data Pre-processing, Data Analysis, and Visualization. The core functionalities of

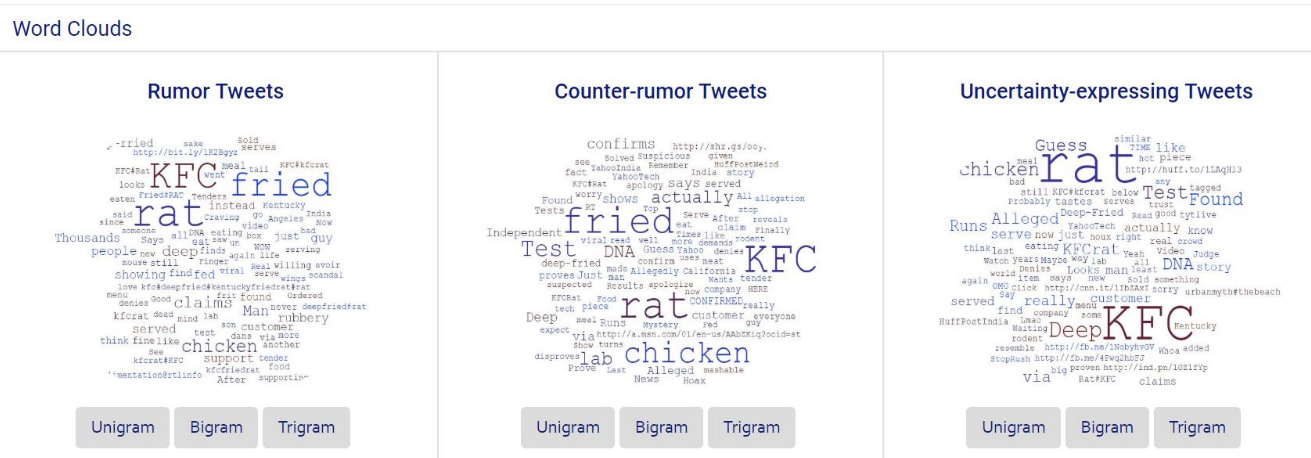

Fig. 4. Word Clouds of the Three Types of Messages.

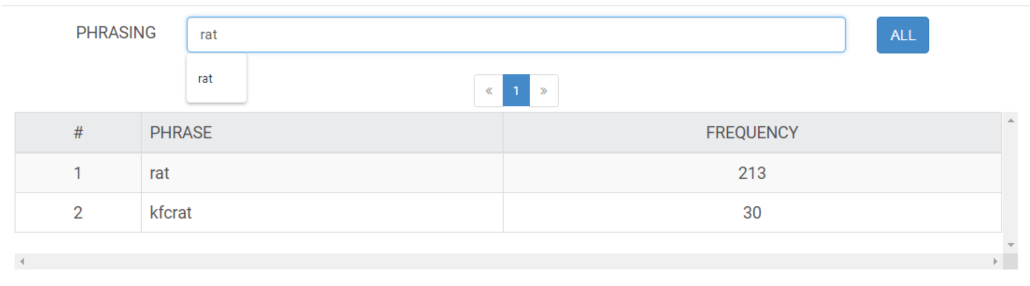

Fig. 5. Search Dashboard to Find Phrases with their Frequency.

ISBN: 978-988-14048-5-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

these modules help to examine the three types of messages—rumor, counter-rumor, and uncertainty-expressing messages—in the wake of a rumor outbreak.

In any rumoring phenomenon, messages become viral easily. While some are false, others bear the truth. Yet other messages raise doubts. The proposed system captures and visualizes these three types of messages in terms of their rise and fall. The combination of the interactive play track and polarized colors gives offer a vantage view as the rumor unfolds

While some recent works have started visualizing rumoring phenomena on social media [30]-[32], future works could invest their efforts to implement an integrated system in which the data stream can be interpreted, analyzed and visualized in real time. With human-in-the-loop process, such a system can support better understandings in making decisions for authorities to deal with online rumors.

References

[1] A. Rojecki and S. Meraz, “Rumors and factitious informational blends: The role of the web in speculative politics,” *New Media & Society*, vol. 18, no. 1, pp. 25-43, 2016.

[2] W. W. Xu and C. Zhang, “Sentiment, richness, authority, and relevance model of information sharing during social Crises—the case of# MH370 tweets,” *Computers in Human Behavior*, vol. 89, pp. 199-206, 2018.

[3] O. Oh, M. Agrawal and H. R. Rao, “Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises,” *MIS Quarterly*, vol. 37, no. 2, pp. 407-426, 2013.

[4] S. Wen, J. Jiang, Y. Xiang, S. Yu, W. Zhou and W. Jia, “To shut them up or to clarify: Restraining the spread of rumors in online social networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 12, pp. 3306-3316.

[5] A. Vista, “Mass media, the ‘sensational message’, and metamorphic truths,” *Telematics and Informatics*, vol. 32, no. 2, pp. 416-423, 2015.

[6] K. Swani, G. Milne and B. P. Brown, “Spreading the word through likes on Facebook: Evaluating the message strategy effectiveness of Fortune 500 companies,” *Journal of Research in Interactive Marketing*, vol. 7, no. 4, pp. 269-294, 2013.

[7] J. Lee and I. B. Hong, “Predicting positive user responses to social media advertising: The roles of emotional appeal, informativeness, and creativity,” *International Journal of Information Management*, vol. 36, no. 3, pp. 360-373, 2016.

[8] P. Ozturk, H. Li and Y. Sakamoto, “Combating rumor spread on social media: The effectiveness of refutation and warning,” in *Proc. of the IEEE Hawaii International Conf. on System Sciences*, New York, 2015, pp. 2406-2414.

[9] A. Pal, A. Y. Chua and D. H. L. Goh, “Does KFC sell rat? Analysis of tweets in the wake of a rumor outbreak,” *Aslib Journal of Information Management*, vol. 69, no. 6, pp. 660-673, 2017.

[10] J. Shin, L. Jian, K. Driscoll and F. Bar, “Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction,” *New Media & Society*, vol. 19, no. 8, pp. 1214-1235, 2017.

[11] K. Starbird, E. Spiro, I. Edwards, K. Zhou, J. Maddock and S. Narasimhan, “Could this be true?: I think so! Expressed uncertainty in online rumoring,” in *Proc. of the CHI Conf. on Human Factors in Computing Systems*, New York, 2016, pp. 360-371.

[12] A. Pal and A. Y. Chua, “Classification of rumors and counter-rumors,” in *Proc. 4th IEEE International Conf. on Information Management*, New York, 2018, pp. 81-85.

[13] A. Friggeri, L. A. Adamic, D. Eckles and J. Cheng, “Rumor cascades,” in *Proc. of the 8th International AAAI Conf. on Weblogs and Social Media*, Palo Alto, 2014, Retrieved from http://static1.squarespace.com/static/540b5d8ae4b08308f411cfc9/t/5414fc87e4b00f3a4f0be566/1410661511886/afriggeri_icwsm2014.pdf

[14] D. Dubois, D. D. Rucker and Z. L. Tormala, “From rumors to facts, and facts to rumors: The role of certainty decay in consumer communications,” *Journal of Marketing Research*, vol., 48, no. 6, pp. 1020-1032, 2011.

[15] Y. Tanaka, Y. Sakamoto and T. Matsuka, “Toward a social-technological system that inactivates false rumors through the critical thinking of crowds,” in *Proc. of the IEEE Hawaii International Conf. on System Sciences*, New York, 2013, pp. 649-658.

[16] J. Ma, D. Li and Z. Tian, “Rumor spreading in online social networks by considering the bipolar social reinforcement,” *Physica A: Statistical Mechanics and its Applications*, vol. 447, pp. 108-115, 2016.

[17] L. Zhao, H. Cui, X. Qiu, X. Wang and J. Wang, “SIR rumor spreading model in the new media age,” *Physica A: Statistical Mechanics and its Applications*, vol., 392, no. 4, pp. 995-1003, 2013.

[18] Q. Zhang, S. Zhang, J. Dong, J. Xiong and X. Cheng, “Automatic detection of rumor on social network,” in *Proc. of the National CCF Conf. on Natural Language Processing and Chinese Computing*, Springer International Publishing, 2015, pp. 113-122.

[19] S. Wang and T. Terano, “Detecting rumor patterns in streaming social media,” in *Proc. of the IEEE International Conf. on Big Data*, 2015, pp. 2709-2715.

[20] N. P. Nguyen, G. Yan, M. T. Thai and S. Eidenbenz, “Containment of misinformation spread in online social networks,” in *Proc. of the 3rd Annual ACM Conf. on Web Science*, Evanston, 2012, pp. 213-222.

[21] A. Y. Chua and S. Banerjee, “Intentions to trust and share online health rumors: An experiment with medical professionals,” *Computers in Human Behavior*, vol. 87, pp. 1-9, 2018

[22] A. J. Kimmel and A. F. Audrain-Pontevia, “Analysis of commercial rumors from the perspective of marketing managers: Rumor prevalence, effects, and control tactics,” *Journal of Marketing Communications*, vol. 16, no. 4, pp. 239-253, 2010.

[23] J. Chae, D. Thom, Y. Jang, S. Kim, T. Ertl and D. S. Ebert, “Public behavior response analysis in disaster events utilizing visual analytics of microblog data,” *Computers & Graphics*, vol. 38, pp. 51-60, 2014.

[24] T. von Landesberger, S. Diel, S. Bremm and D. W. Fellner, “Visual analysis of contagion in networks,” *Information Visualization*, vol. 14, no. 2, pp. 93-110, 2015.

[25] K. Toutanova, D. Klein, C. D. Manning, and Y. Singer, “Feature-rich part-of-speech tagging with a cyclic dependency network,” in *Proc. of the Conf. of the Association for Computational Linguistics on Human Language Technology-Volume 1*, 2003, pp. 173-180.

[26] A. P. Rauth and A. Pal, “Design and evaluation of Text Pre-Processor: A tool for text pre-processing,” *AMSE Journals IIETA Publication-2017-Series: Advances A*, vol. 54, no. 2, pp 169-184, 2017.

[27] A. Y. Chua, & S. Banerjee, “A Study of Tweet Veracity to Separate Rumours from Counter-Rumours,” in *Proc. of the 8th International ACM Conf. on Social Media & Society*, New York, 2017, p. 4.

[28] J. W. Pennebaker, C. K. Chung, M. Ireland, A. Gonzales, and R. J. Booth, “The development and psychometric properties of LIWC 2007,” Austin, TX: LIWC.net, 2007.

[29] O. Oh, K. H. Kwon and H. R. Rao, “An exploration of social media in extreme events: rumor theory and Twitter during the Haiti Earthquake 2010,” in *Proc. of the International Conf. on Information Systems*, St. Louis, 2010, Retrieved March 12, 2016 from http://aisel.aisnet.org/cgiviewcontent.cgi?article=1223&context=icis2010_submissions.

[30] A. Dant, and J. Richards, “Behind the rumours: How we built our Twitter riots interactive,” 2011. Retrieved Sep 5, 2014 from https://www.theguardian.com/news/datablog/2011/dec/08/twitter-riots-interactive

[31] A. Dang, A. Moh’d, E. Milios and R. Minghim, “What is in a rumour: Combined visual analysis of rumour flow and user activity,” in *Proc. of the 33rd ACM International Conf. Computer Graphics*, 2016, pp. 17-20.

[32] C. Shao, G. L. Ciampaglia, A. Flammini and F. Menczer, “Hoaxy: A platform for tracking online misinformation,” in *Proc. of the 25th International Conf. Companion on World Wide Web*, 2016, pp. 745-750.

ISBN: 978-988-14048-5-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)