A Study on Rumor Propagation Trends and Features in a Post Disaster Situation

Tamal Mondal†, Tunir Roy, Indrajit Bhattacharya
Department of Computer Application
Kalyani Govt. Engg. College, India
{tamalkalyanigov, tunir.27, indra51276}@gmail.com

Sourav Bhattacharya, Indranil Das
Department of Information Technology
Govt. College of Engg. & Textile Technology Serampore, India
{souravbhattacharya1997, indranildas213}@gmail.com

ABSTRACT

This paper explores the characteristics of three established rumor propagation features i.e. *temporal, structural and linguistic* for collected tweets in a natural disaster i.e. Chennai Flood, 2015. The consequences of the features and the rationale behind such effects have been explored extensively for collected rumors and non-rumors of that disaster event.

CCS CONCEPTS

• Information Systems~ Collaborative filtering • Information Systems~ Social networking sites • Information Systems~ Data stream mining

KEYWORDS

Rumor, Features, Disaster, Propagation, Twitter

†Corresponding Author

ICDCN '19, January 4–7, 2019, Bangalore, India

ACM ISBN 978-1-4503-6094-4/19/01..$15.00
https://doi.org/10.1145/3288599.3295581

1 Background & Contribution

Rumor detection in post disaster scenario has become a trending topic of research in past few years. Although, the theories on rumors and non-rumors have been proposed over many decades. In some existing literatures [1, 2, 3], rumors are generally termed as "*Unverified or false factual statement.*" In order to wipe out inconsistencies, detection and mitigation of such rumors in critical time and controlling their propagation in online social media (*like twitter*) should be the primary concern in post disaster scenario. Yet, ahead of that, the characterization of rumor features in contrast to rumors and non-rumors are necessary. In past literatures [4, 5], the rumors have been featured on various aspects *i.e. propagation time, propagators, structural contents etc.* to portray a comprehensible understanding. From these studies, eminent rumor propagation features can be easily outlined as **Temporal, Structural and Linguistic Feature** [4, 5]. Now, based on such rumor propagation aspects several popular rumor analysis and detection models have been proposed in recent past [6, 3, 7]. Most of these existing models captured data from various disaster types like Boston bombing, Chennai Flood etc. to dectectrumors using the outlined aspects mentioned. However, the question lies on whether the attitude of such aspects changes on the nature of the disaster? In order to interpret such natures of the rumor propagation features, the behavioral patterns of each feature on rumors and non-rumors should be monitored closely. Furthermore, the reason behind such behaviors and acceptability of the features for distinguishing rumors and non-rumors should also be the concern. Considering such key phenomenon, in proposed approach we have evaluated 3 rumor propagation aspects (temporal, structural and linguistic) for rumor and non-rumor event set, collected from twitter during Chennai Flood, 2015.The consequences of each aspect on those rumor and non-rumor events have been clearly represented through vivid and explicit details. In inclusion to that, the rationale behind such effects have been discussed and illustrated through proper justifications. Note, the proposed approach can be distinguishable from the literatures discussed in [7, 8, 11]. Unlike the work in [7, 8], the current work critically explores not only the behavioral pattern but also the rationale behind such behaviors of rumor propagation features under the shade of rumor and non-rumor topics. Besides, in [11], a rumor propagation model has been designed through utilizing the rumor propagation aspects in order to detect rumors at early stage. However, it can be observed that the model suffers from false positivity. That is, it wrongly predicts around 800 tweets as rumors. Therefore, for efficient utilization of rumor propagation aspects in identifying rumors, one should effectively study the behavioral framework of each aspect. The distinguishing characteristics of rumors and non-rumors under the shade of those aspects should also become understandable. In proposed literature, such an attempt has been made. Through a case study from collected tweets regarding different rumor and non-rumor events of Chennai Flood, we not only have tried to make out the effects of rumor propagation features but also aimed to establish the rationale behind such effects. Furthermore, based on the characteristics of each aspect over rumor and non-rumor events, the suitability in differentiation of rumors and non-rumors has been discussed.

2 Identification & Validation of Rumors and Non-rumors

We have collected approximately 1.194 millions of tweets related to Chennai Flood, 2015 that comprise the profile information of 2.61 lakh distinct twitter users. The tweets have been collected in between December 1, 2015 – December 10, 2015 when distinct parts of Chennai city were severely affected by heavy rainfall. Now, in order to identify rumor and non-rumor events we have considered popular websites of renowned news media like NDTV, Deccan Chronicle, The Hindu etc. We have selected four topics (2 rumors and 2 non-rumors) those became popular through circulation in twitter during the disaster. The selected rumor and non-rumor topics are Crocodile Escape (Rumor), Chembarambakkam Lake Overflow (Rumor), Little Girl Lost (Non-Rumor) and Passport Lost (Non-Rumor) respectively. Now, in order to group tweets related to four topics (mentioned above), we have used K-means clustering approach. With the value of K =10, we have specified the number of clusters that will be formed on collected set of 1.194 million tweets. Tweets that are more likely to each other are grouped together in order to form strict hard clusters. From these newly generated clusters, we find the labels that are associated with them. The cluster labels give us an indication of what types of content words are frequently occurring in tweets that are present in each cluster. Now, based on the rumor and non-rumor topics selected, we find the closest labels that match with those topics. If a match has been found, that means the cluster is populated with tweets of that particular rumor or non-rumor topic. Then the tweets related to other topics that have less influence over the cluster are simply discarded and the cluster has been considered rumor or non-rumor cluster that are populated with tweets related to mentioned rumor or non-rumor event. In such a manner, we have identified four event clusters related to two rumor and two non-rumor topics mentioned above. In order to ensure that the tweets belonging to each rumor and non-rumor clusters are valid, we asked two participants to analyze each of the tweet clusters. They also had been provided the links of different websites of news media where the original rumor and non-rumor topics have been posted. This process yielded 1) around 703 and 700 rumor tweets (including retweets) related to two rumor events i.e. Crocodile Escape and Chembarambakkam Lake Overflow and 2) around 6456 and 15350 non-rumor tweets (including retweets) related to two non-rumor topics i.e. Little Girl Lost and Passport Lost respectively. The spearman"s correlation coefficient that measures the amount of agreement among participants was 0.90.

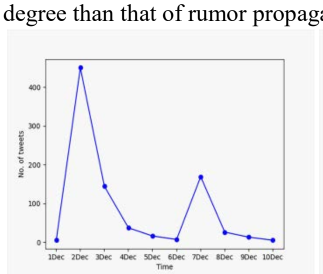
3 Rumor Propagation Feature Analysis

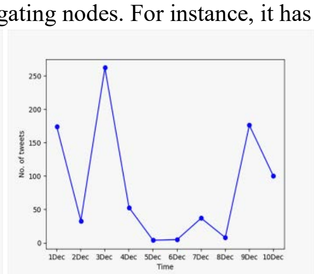
3.1 Temporal Aspect

Temporal feature of a rumor, as the name suggests, is the effect of time in the propagation of a rumor. Here, we try to characterize the distribution of the rumor and non-rumor tweets regarding an event in accordance to the tweet generation timestamp. We then visualize the event by plotting a curve of the frequency of such tweets to the date it has been posted. It has been observed that rumors have recurring distinctive spikes or peaks, whereas non-rumors have only a single spike (Figure 1). This can be attributed to the fact that rumors as a whole have both the highest diffusion power and incurs the most friction in its propagation. Due to the lack of proper news media interventions, rumor propagators often place statements that are emotive or delicate to the community. Such controversial statements are either believed or questioned by the community due to absence of proper evidences. Therefore, any rumor topic frames apprehension among the people and they desire to know various facts regarding that event. From the psychological point of view, rumors always warrant a discussion whether it is believed or not. In addition, rumor inherently morphs itself, thus a new discussion starts in accordance with its belief and disbelief. Non-rumors on the other hand, may or may not have a higher diffusion power but it incurs the least friction, as non-rumors are more neutral in nature and thus it has a lower chance of people disbelieving it. Note, unlike non-rumors, rumors have a tendency to stimulate people towards its content. In disaster situation, such attraction increases rapidly due to information uncertainty. That leads rumorous events to propagate and to become hot topic for discussion. Table 1 depicts how the rumor events became hot topics for discussion in twitter from December 1 – December 10, 2015. It can be observed that, regarding the rumor topic of crocodile escape, it takes around 14 hours to identify that the topic was a rumor (Table 1). But, in the duration of 14 hours the topic became so much viral in twitter that many of the people actually believed the topic as genuine one. As a result, though the first disbelief statement on crocodile escape posted on '2015-12-02 04:56:42" and numerous posts have been made from different handles to make people aware about the rumor, but still people were expressing their skepticism and belief about the topic (Table 1). In the same way, for the rumor topic of chembarambakkam lake, it takes around 86 hours (approx.) to identify the topic as rumor. Due to that high diffusion power, the rumor persisted in the network for a prolonged duration. These two scenarios on rumor topics clearly elucidate the reason behind such recurring distinctive peaks illustrated in Figure 1 (a) & (b). The deficiency of information has made different users believe, disbelieve or question the rumor topic. These activities transform the rumor topic to a focal point for public discussion. Consequently, people find interests on various facts regarding the topic at different phase, which keep the topic alive for a longer period. On the other hand, it can be observed from Figure 1(c) & (d) that non-rumor topics become popular for a particular period and after that is often found insignificant by the community.

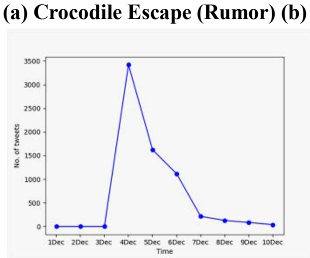
3.2 Structural Aspect

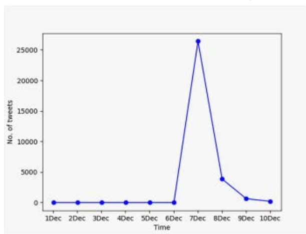
In structural feature, we study how a rumor or non-rumor propagates in a closed network, the closed network being the followers of the person who originally initiated the tweets. In a retweet, the tweet has a source and a target, the source being the person who originally posted the tweet and the target being the the person who actually retweeted it. From our annotated data in each rumor and non-rumor topic cluster, we visualize the findings in the form of a connected graph (Figure 2), where the nodes are the people who posted or retweeted the rumor or non-rumor tweet and the edges being the source target relationship of each retweet. We can observe that rumor topics tend to have greater number of source nodes than non-rumors. This is due to the fact that people can change the rumor by morphing the original tweet to a new tweet and thus increasing the propagation of the rumor. In non-rumors the source nodes are lesser than that of rumors as people tend to generally retweet a non-rumor more frequently than morphing the content of the tweet and then posting it. Since rumors generally have a larger friction than non-rumors, we observed that rumor tweets were significantly less than non-rumors tweets. Also from these graphs (Figure 2) we can find out the edge with the highest degree, this being the node which has attributed to the highest propagation of the rumor or non-rumor in its respective cluster. Furthermore, it can also be inferred that the nodes, which propagate non-rumors, have significantly higher degree than that of rumor propagating nodes. For instance, it has





(a) Crocodile Escape (Rumor) (b) Chembarambakkam Lake (Rumor)





(c) Little Girl Lost (Non-Rumor) (d) Passport Lost

Figure 1: Time Series Representation of Rumors and Non-rumors.

Table 1: Snapshot of the discussion on Rumor topics in twitter

Crocodile Escape (Rumor)
given the flooding levels in chennai does anyone know what is happening in the crocodile bank? #chennairains', **'2015-12-01 11:21:41**
20 crocodiles missing from zoo in chennai', '2015-12-01 17:29:58
be alert chennai peoples 20 crocodiles escaped from crocodile park.. #chennaifloods #chennairainshelp https://t.co/l3y3vmhfct', **'2015-12-02 04:56:42**
apparently crocodiles are swimming in the floods in chennai', '2015-12-03 08:21:15
so rumours. the fine city of chennai did not face a crocodile apocalypse nor did it face some sort of supercyclone. it did face politics.', '2015-12-10 04:04:48
nasa picture shows chennai boy feeding tiger biscuits to escaped crocodiles. c'mon whatsapp user viral this", '2015-12-10 04:11:06
Chembarambakkam Lake Overflow (Rumor)
chennai collector issues flood alert warning; adyar will swell as outflow from chembarambakkam is now 5 000 cusecs #chennairains', **'2015-12-01 06:02:08**
chembarambakkam river water discharge increased to 30 000 cusecs today..it was 5000 yesterday.god please save my chennai..#prayforchennai', '2015-12-01 17:24:47
'chembarambakkam dam opened. flooding in ekaatuthaangal adyar and other areas. join hands and help chennai _/_ #chennaifloods', '2015-12-02 01:38:54
please do not believe flood rumors in chennai. as of now chembarambakkam porur safe. no water release. stop spreading panic.', **'2015-12-04 20:13:28**
chennai floods: what happened at chembarambakkam negligence or nature"s fury? https://t.co/lxstpmlsdw https://t.co/zgtrkaiqcr', '2015-12-10 00:50:59

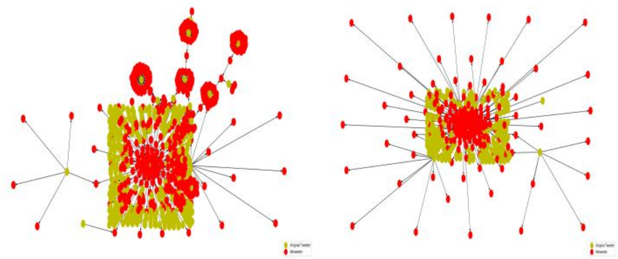
been outlined in Table 2 that for rumor topic crocodiles escape the user handle that had the highest degree was @ 'Sibi_Sathyaraj' with the degree of 97. On the other hand, for rumor topic Chembarambakkam lake the user handle that had the highest degree was @„ChennaiConnect' with the degree of 121. This means that, these two handles were the highest influencing nodes that had influenced the maximum number of users through their rumor related posts. Nevertheless, in terms of non-rumor topics, the influencing capabilities of highest degree nodes are far more than that of rumors (Table 2). The highest influencing nodes for non-rumor topics little girl lost and passport lost were @ 'shrutihaasan' and @'SushmaSwaraj' respectively (Table 2). In addition to the differentiation between rumors and non-rumors over the number of influencing nodes, it has been depicted in Section 3.2 that rumors have more number of originating nodes than that of non-rumors. This observation also has been clearly shown in Table 3. It can be observed that, though rumors have large number of originating nodes, the total number retweets have been found to be significantly lesser compared to non-rumors. These insights clearly show that people often morph rumors rather retweeting it. The information uncertainty regarding rumor topics in post disaster scenario often leaves a ground for discussion among the twitter community about the validity of the topic. Therefore, instead of simply retweeting the fact, most of the people contrive the news through verification or correction statements and propagate it further into the twitter network.

3.3 Linguistic Aspect

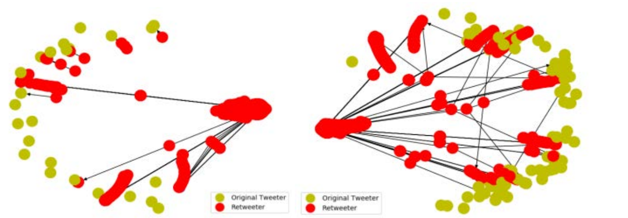
In linguistic feature, we study the inherent nature of the rumor and non-rumor topics. Linguistic features generally focuses on the tweet"s structure i.e. it"s *sentiment and controversial* level. In previous literatures [7, 8], it has been mentioned that unlike non-rumors, rumors are dominated by *high sentiments, uncertainty or controversy*. In order to evaluate such conception, we have developed two modules to analyze the impact of linguistic aspect on rumors and non-rumors in terms of *polarity of sentiment* and *controversy*.

3.3.1. Sentiment Analysis: In order to calculate the sentiment score, we have used the *TextBlob package in Python*. *TextBlob* returns two properties i.e. average *polarity and subjectivity* of any statement. Polarity is a fraction that lies in the range of [-1, 1] where 1 means extreme positive statement, -1 means a extreme negative statement and 0 means neutral statement. We have annotated rumor and non-rumor events through a) strict hard clustering and b) discarding those tweets that are non- subjective to the events (Section 2). Hence, we have only evaluated sentiment polarity for rumor and non-rumor events. For any text, depending on the sense of each content word, the corresponding polarity value is returned. After that, the mean of such polarities is rated which would then be the sentiment polarity for that particular text. The polarity score can be > 0, < 0 and = 0 indicating positive, negative and neutral sentiment respectively. Now, the sentiment polarity score of each tweet regarding rumor and non-rumor event cluster has been evaluated and illustrated through Figure 3. For every rumor and non-rumor event cluster, we convert the discrete sentiment scores of the tweets to a continuous curve. It can be observed that, in the case of rumor events i.e. *crocodile escape* and *Chembarambakkam Lake,* the graphs represent non-uniform nature with varying positive, negative and neutral polarity for increasing number of tweets (Figure 3 (a) & (b)). In contrast, for non-rumor events i.e. *Little Girl Lost* and *Passport Lost*, the nature of the graphs are uniform and reside in the neutral polarity (Figure 3 (c) & (d)). The rationale behind such representations can be attributed to fact that, rumor events have more number of originating nodes thus they have a higher impact of propagation. Besides, during the propagation of rumor events, they face huge amount of criticism among the community. Unlike non-rumors, people often express their sentiments like *anxiety, anger, sorrow etc.* regarding rumor topics. In contrary, for non-rumor events, the public sentiments are more neutral than that of rumors. The lower number of originating nodes clearly indicates that people often prefer retweeting the topic rather arguing in case of non-rumor events.

3.3.2. Controversy Analysis: The formulation for evaluation of tweet controversy has been used in [11]. In proposed work, we are using the identical approach with a slight conceptual revision.



(a) Crocodile Escape (Rumor) (b) Chembarambakkam Lake



(c) Little Girl Lost (Non-rumor) (d) Passport Lost (Non-rumor)

Figure 2: Structural Representation of Tweets regarding Rumor and Non-rumor Events.

Table 2: Highest degree nodes for rumor and non-rumor events

Topic	Highest Degree Handle	Influenced Nodes
Crocodile Escape (Rumor)	@'Sibi_Sathyaraj'	97
Chembarambakkam lake (Rumor)	@,ChennaiConnect'	121
little girl lost (Non- rumor)	@'shrutihaasan'	6000
passport lost (Non-rumor)	@'SushmaSwaraj'	13104

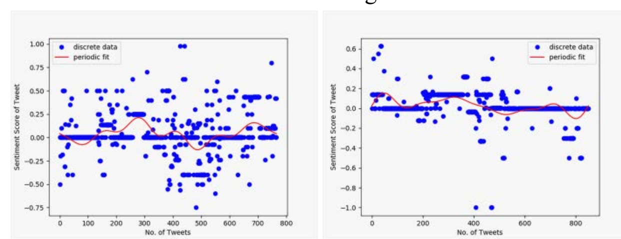
Table 3: Total number of originator and retweets for rumor and non-rumor events

Topic	Originator	Retweet
Crocodile Escape (Rumor)	301	402
Chembarambakkam lake (Rumor)	220	480
little girl lost (Non- rumor)	32	6424
passport lost (Non-rumor)	86	15264

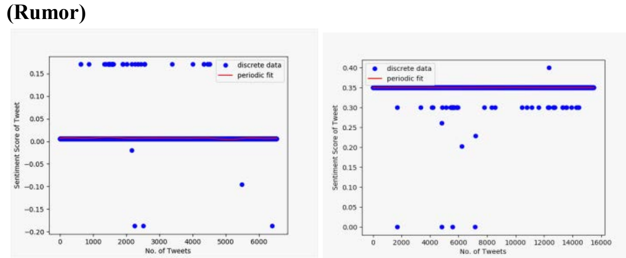
Instead of calculating the controversial score for each tweet, we are interested in the evaluation of overall controversy score of each rumor and non-rumor event. In order to evaluate the controversial score for any rumor or non-rumor topic, a) we have calculated sentiment score for each tweet belonging to the topic, b) sorted the tweets based on the type of sentiments *(positive/negative)* they associated with, c) evaluated the mean positive or negative polarity *(*$|Pos|$ *or* $|Neg|$) score by adding the positive or negative scores associated with the respective tweets and then dividing by the number of tweets of that particular polarity type and d) The value of $|Neu|$ has been obtained using the formula: $(1- (|Pos| + |Neg|))$ (*as polarity value lies in [-1, 1]* (Section 3.3.1)). The method for evaluation of polarity score for each tweet has already been discussed in Section 3.3.1. Now, the controversial score is calculated using the formula in [11]. Table 4 depicts the controversial scores for rumor and non-rumor topics.

4 Conclusion

The proposed work presents a deep insight about the behavioral patterns of three rumor propagation features on collected rumor and non-rumor events of Chennai Flood, 2015. Differentiation between rumors and non-rumors has been clearly depicted under the shed of these rumor propagation features with illustrations. The current work has analyzed the natures of rumors and non-rumors and the background of such natures. However, rigorous analysis about the consistency of such behaviors based on disaster types and disaster locations is still untouched. Besides, the essence of rumor propagation aspects over early stage rumor detection has been remained a burning issue.



(a) Crocodile Escape(Rumor) (b) Chembarambakkam Lake (Rumor)



(c) Little Girl Lost (Non-rumor) (d) Passport Lost (Non-rumor)

Figure 3: Linguistic Aspects of Rumor and Non-rumor Events

Table 4: Controversial scores of rumor and non-rumor topics

Topic	Controversy
Crocodile Escape	0.510844
Chem Lake	0.231081
Little Girl	0.007551
Passport	0.00

REFERENCES

[1] Bordia, P. and DiFonzo, N., 2004. Problem solving in social interactions on the Internet: Rumor as social cognition. Social Psychology Quarterly, 67(1), pp.33-49.

[2] Buckner, H. T. (1965). A theory of rumor transmission. Public Opinion Quarterly, 29(1), 54–70.

[3] Zhao, Zhe, Paul Resnick, and Qiaozhu 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.

[4] Kwon, Sejeong, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. "Aspects of rumor spreading on a microblog network." In International Conference on Social Informatics, pp. 299-308. Springer, Cham, 2013.

[5] Kwon, Sejeong, Meeyoung Cha, Kyomin Jung, Wei Chen, and Yajun Wang. "Prominent features of rumor propagation in online social media." In 2013 IEEE 13th International Conference on Data Mining, pp. 1103-1108. IEEE.

[6] Zhang, Qiao, Shuiyuan Zhang, Jian Dong, Jinhua Xiong, and Xueqi Cheng. "Automatic detection of rumor on social network." In Natural Language Processing and Chinese Computing, pp. 113-122. Springer, Cham, 2015.

[7] Mondal, Tamal, Prithviraj Pramanik, Indrajit Bhattacharya, Naiwrita Boral, and Saptarshi Ghosh. "Analysis and Early Detection of Rumors in a Post Disaster Scenario." Information Systems Frontiers (2018): 1-19.