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

Tamal Mondal[1], Tunir Roy[1], Indrajit Bhattacharya[1], Sourav Bhattacharya[2], Indranil Das[2]
[1]Department of Computer Application, Kalyani Government Engineering College, Kalyani, India.
[2]Department of Information Technology, Government College of Engineering & Textile Technology Serampore, India.
{tamalkalyanigov, tunir.27, indra51276, souravbhattacharya1997, indranildas213}@gmail.com

ABSTRACT

In today's world, mitigating the effect of rumor, which occurs in accordance to any prominent event, is of paramount importance. In order to reduce such effect, we need to analyze the trends and patterns that help in rumor propagation in online social media *like Twitter*. In this paper, we have evaluated the characteristics of three established rumor propagation features i.e. *temporal, structural and linguistic* 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. Moreover, through such interpretation we want to infer, a) the behavioral patterns of rumor propagation features for a particular disaster event b) whether these features are capable of differentiating the nature of rumors and non-rumors.

KEYWORDS

Rumor, Features, Disaster, Propagation, Twitter

1 Background & Motivation

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 [1, 2, 3, 4]. In some existing literatures [4, 5, 6], rumors are generally termed as "*Unverified or false factual statement.*" That means, any statement regarding any event that is unverified or false, can considered as rumors. Effective identification 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. In order to wipe out inconsistencies from the generated information in online social media (*i.e. twitter*), recognition and mitigation of rumors are required. Yet, ahead of that, the characterization of rumor features in contrast to non-rumors is necessary. In past literatures [7, 8], the authors have featured rumors on various aspects *i.e. propagation time, propagators, structural contents etc.* to portray a comprehensible understanding about rumors. From these studies, eminent rumor propagation features can be easily outlined as follows,

Temporal Feature: The lifetime of rumors is very short as they use that fertile ground where the information uncertainty takes place. The requirement of proper information about any event in absence of any authentic and verifiable sources often creates hoaxes. As a result, statements related to verifications, corrections, interrogations etc. regarding any event / topic become rumors [8].

Structural Feature: Rumors spread rapidly in a denser network as a gossip as the rumor messengers and the receivers of those rumors are related to each other in online social media. As a result, the rumors spread more widely in sparse structures [8].

Linguistic Feature: Rumors are dominated by high sentiments, controversies or uncertainties [7].

Now, based on such rumor propagation aspects several popular rumor analysis and detection models have been proposed in recent past [9, 10, 6, 11]. In [9], authors have proposed a rumor detection approach based on sub-categorized features like sentiment and opinion polarity, social influence, popularity orientation etc. along with some other features like rumor content features and blogger features. The authors in [10], captured the time series modeling of rumor's life cycle. In [6], authors proposed a model through tracing and clustering of skeptic statements, identifying similar texts that does not contain those skeptic phrases and rank each of those clusters. In [11], a probabilistic rumor detection model has been developed by combining the prominent features and content-based analysis of tweets in order to evaluate the tweet statements those are highly probable of being rumors. Many of these existing models captured data from different disaster types *like Boston bombing*[1]*, Chennai Flood*[2] *etc.* to detect rumors using the outlined aspects (*mentioned above*). But, the question lies on whether the attitude of such aspects changes based on the nature of the disaster? In order to investigate such issue, the behavioral patterns of rumor propagation features and the rationale behind such behavior should be explored effectively. Besides, the contributions of the aspects in distinguishing rumors and non-rumors should also be analyzed. In various past literatures [7, 8, 11], the authors have tried to feature rumors and render a coherent insight about rumors and non-rumors for different topics including disaster events. But, the minute experimentations of rumor propagation features and detailed explanations about the cause of behavioral patterns of these features on rumors and non-rumors in any disaster event have not been performed yet. Therefore, in precise we are seeking for the answers of following questions,

a) *What would be the behavioral patterns of rumor propagation features on rumors and non-rumors?*
b) *What is the rationale behind such behaviors?*

[1]https://en.wikipedia.org/wiki/Boston_Marathon_bombing
[2]https://en.wikipedia.org/wiki/2015_South_Indian_floods

2 Contribution

In this Section, the contribution of the proposed literature has been discussed. It has been depicted in Section 1 that in order to interpret the nature of 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, 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, for effective 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 have not only 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. The rest of the paper has been organized as follows,

In Section 3, the identification of rumor and non-rumor events and testing of the validation of rumors and non-rumors regarding the events in Chennai Flood have been performed. The behavioral patterns of each three aspect (*temporal, structural and linguistic*) for the collected rumor and non-rumor events have been analyzed in Section 4. The reason behind such characteristics and a clear distinction of rumors and non-rumors under those characteristics been discussed with proper justifications. Finally, we conclude the paper in Section 5.

3 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[3]. Now, in order to identify rumor and non-rumor events we have considered popular websites of renowned news media like *NDTV*[4], *Deccan Chronicle*[5], *The Hindu*[6] *etc.* We have selected four topics (*2 rumors and 2 non-rumors*) those became popular through circulation in twitter during the disaster. A brief introduction of those rumor and non-rumor events is as follows,

Crocodile Escape (Rumor): There was circulating news about crocodiles escaping from the Madras Crocodile Bank in Mamallapuram. But later, the crocodile bank confirmed that it was a rumor.

Chembarambakkam Lake Overflow (Rumor): There was a news circulation in social media that the Chembarambakkam Lake is full and another flood disaster is coming. However, later it has been found as rumorous topic.

Little Girl Lost (Non-Rumor): There was a news event about a little girl who has been lost during Chennai Flood. This news circulated in twitter for several days.

Passport Lost (Non-Rumor): A news event posted by external affairs minister of India on passport issue states that people of Chennai who lost their passport documents during flood will be reissued passport with free of charge.

Now, in order to group tweets related to four topics (mentioned above), we have used K-means clustering approach [12]. 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 type 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. As an example the label might be *'crocodile' or 'crocodile escape'* and our rumor may be of the form *"40 crocodiles escaped during Chennai floods"*. 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 belong to each rumor and non-rumor clusters are valid, we asked two participants to analyze each tweet of rumor and non-rumor cluster. We also had provided them the links of different websites of news media where the original rumor and non-rumor topics have been posted. This process yields the number of tweets that are actually related to any of the events are as follows,

Crocodile Escape: Approx. 703 rumor tweets (*including retweets*) have been found valid and related to this particular event.

Chembarambakkam Lake Overflow: Approx. 700 rumor tweets (*including retweets*) have been found valid and related to this particular event.

Little Girl Lost: Around 6456 non-rumor tweets (*including retweets*) have been found valid and related to this particular event.

Passport Lost: Around 15350 non-rumor tweets (*including retweets*) have been found valid and related to this particular event.

The spareman correlation coefficient [13] that measures the amount of agreement among participants was 0.90. Some examples of tweets related to four events have been listed in Table 1.

[3]https://en.wikipedia.org/wiki/2015_South_Indian_floods
[4]https://www.ndtv.com/chennai-news/in-flooded-chennai-crocodile-escaped-rumours-are-denied-1250207
[5]https://www.deccanchronicle.com/nation/in-other-news/031117/stop-flood-of-rumours.html
[6]https://www.thehindu.com/news/cities/chennai/chennai-floods-rumours-rain-down-on-online-platforms/article7953661.ece

Table 1: Examples of tweet related to four topics

Topic Name	Tweets Examples
Crocodile Escape (Rumor)	20 crocodiles missing from zoo in Chennai.
	given the flooding levels in chennai does anyone know what is happening in the crocodile bank? #chennairains
Chembarambakkam Lake Overflow (Rumor)	flood warning along adyar in #chennai; 20000 cfps of water from chembarambakkam reservoir released.
	#chennai #rain alert : water from #chembarambakkam lake may overflow.
Little Girl Lost (Non-Rumor)	this little girl is lost in chennai floods pls help her find her parents. pls share as much as you can.
	this small girl is lost in chennai floods pls help her to meet her parents. dcp sivakumar
Passport Lost (Non-Rumor)	if your passport is lost or damaged in floods pl go to any of three psks in chennai. they will issue u fresh passport free of charge. pl rt
	rt @sushmaswaraj if ur passport is lost or damaged in floods pl go2 any of 3 psks in #chennai. will issue u fresh passport free of charge.

4 Rumor Propagation Feature Analysis

In this Section, the three rumor propagation features i.e. *temporal, structural and linguistic* are analyzed in the context of four event topics of Chennai Flood, 2015. Here, the behavioral patterns of each aspect have been monitored firmly and reasons behind such behaviors have been documented well with proper illustrations.

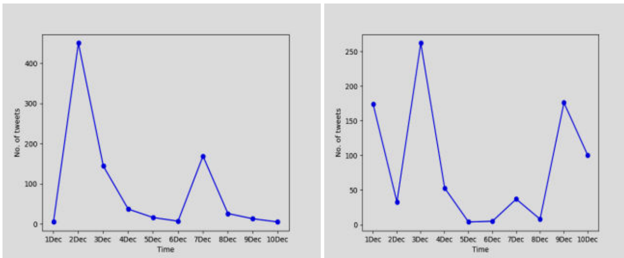
4.1 Temporal Aspect

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 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 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. Also, rumors 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 2 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 2). 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 that rumor, but still people were expressing their skepticism and belief about the topic (Table 2). 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.

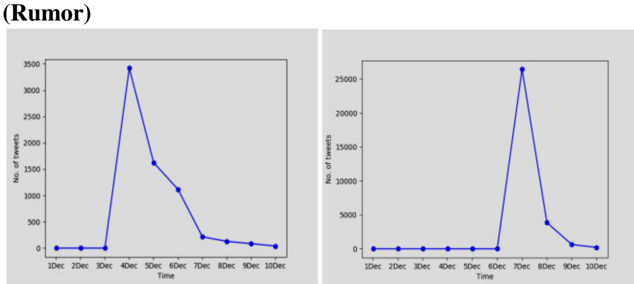
Table 2: Snapshot of the discussion on Rumor topics in twitter

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 chennai friends pls stop doing this fake news !!! crocodiles in d city and all other nonsense are just rumours.... https://t.co/sxaiyuhv9w', '2015-12-02 06:53:17 in flooded chennai 'crocodiles have escaped' rumours are denied https://t.co/s2eugviu6v via @ndtv", '2015-12-03 00:02:38 so 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 @reks612 no mobile network/landlines working in chennai. please try for some other way of communication once you cross chembarambakkam lake', '2015-12-03 04:43:05 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

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 twitter community.



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

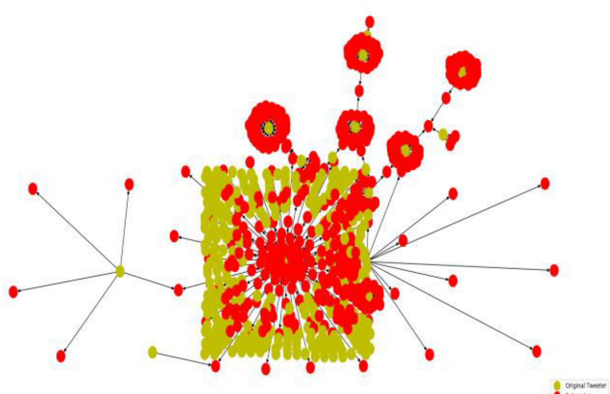


(c) Little Girl Lost (Non-Rumor) (d) Passport Lost

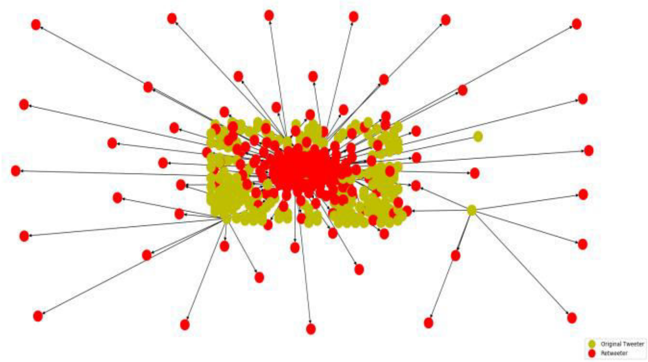
Figure 1: Time Series Representation of Tweets regarding Rumor and Non-rumor Events

4.2 Structural Aspect

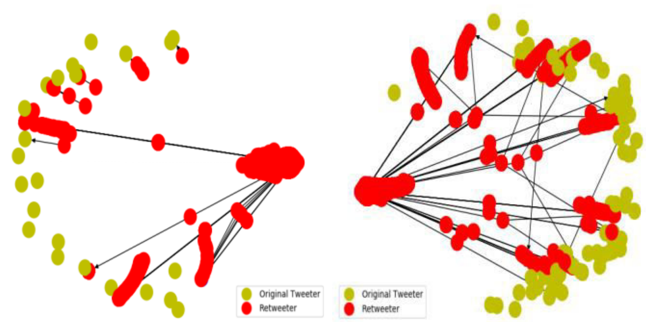
In structural feature, we study how a rumor or non-rumor propagates in a closed network, the closed network being the followers of the person who originally initiated the tweets. In a retweet, the tweet has a *source* and a *target*, the *source* being the person who originally posted the tweet and the *target* being the 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 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 been outlined in Table 3 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*



(a) Crocodile Escape (Rumor)



(b) Chembarambakkam Lake Overflow (Rumor)



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

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

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 user handles 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 3). The highest influencing nodes for non-rumor topics *little girl lost* and *passport lost* were @ *'shrutihaasan'* and @*'SushmaSwaraj'* respectively (Table 3). In addition to the differentiation between rumors and non-rumors over the number of influencing nodes, it has been depicted in Section 4.1 that rumors have more number of originating nodes than that of non-rumors. This observation also has been clearly shown in Table 4. 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.

Table 3: 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

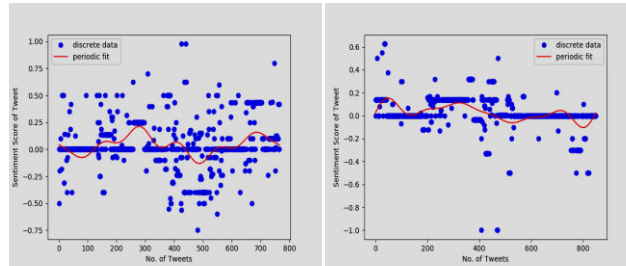
Table 4: Total number of originator and retweets for rumor and non-rumor events

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

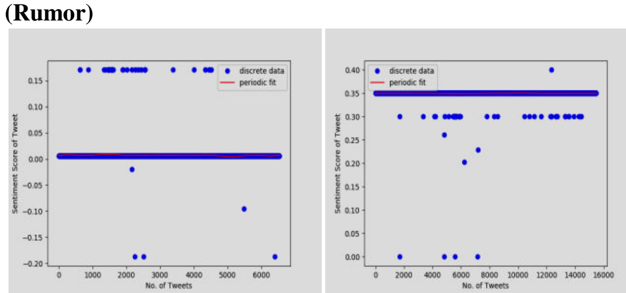
4.3 Linguistic Aspect

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*.

4.3.1. Sentiment Analysis: In this Section, the sentiment polarity of each tweet regarding rumor and non-rumor topics is calculated. In order to calculate the sentiment score, we have used the *TextBlob package in Python* [14]. The *TextBlob package* for Python is a convenient way to do lot of NLP tasks (*POS tagging, Sentiment analysis, Classification etc.*). *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. Subjective sentences generally refer to *personal opinion or judgment* that is also a fraction that lies in the range of [0, 1]. 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 3). Hence, we have evaluated sentiment polarity for rumor and non-rumor events. We have used the method *TextBlob(text as parameter).sentiment.polarity* that returns the polarity value for any text. The method utilizes *en-sentiment.xml* [14] file where different types of senses, polarity values etc. for millions of content words and their synsets have been provided. For any text, depending upon the sense of each content word, the corresponding polarity value will be returned. After that, the mean of such polarities is evaluated which would then be the sentiment polarity for that particular text. The polarity score can be > 0, < 0



(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

and = 0 indicating positive, negative and neutral sentiment respectively. The polarity score of some sample tweets have been depicted in Table 5.

Table 5: Sentiment Polarity score of some tweets.

the #crocodile escaping story of #chennai being circulated on whatsapp is false .	-0.400 (Negative)
#chennaifloods:breach of #porurlake feared; rumours of #crocodiles #zooanimals on the loose are false	-0.238 (Negative)
just to scotch any mischievous rumours about #chennai crocodile bank. all is fine there"	0.416 (Positive)

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 representing 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. Therefore, we can state that, non-rumor topics have lesser impact on the community.

4.3.2. Controversy Analysis: It has been discussed in [11] that controversial event provokes a public discussion where people express their disbelief or opinion about the event. Controversial tweets often instigate people to express their views, doubt about the fact mentioned in the tweet and this may lead to rumors in future. In order to investigate that, we have evaluated the overall *controversial score* of each rumor and non-rumor event. The formula for evaluation of tweet controversy has been used earlier in [11]. In proposed work, we are using the identical approach with a slight conceptual modification. 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. The given formula for evaluating controversial score is as follows,

$$\boldsymbol{Controversy\_Score} = \frac{Min(|Pos|,|Neg|)}{Max(|Pos|,|Neg|)} \times \frac{|Pos|+|Neg|}{|Pos|+|Neg|+|Neu|} \quad (1)$$

Where,

|Pos| = Mean positive sentiment associated with the rumor or non-rumor topic.

|Neg| = Mean negative sentiment associated with the rumor or non-rumor topic.

|Neu| = Mean neutral polarity associated with the rumor and non-rumor topic which is (1- (|Pos| + |Neg|)).

In order to evaluate controversial score for any rumor or non-rumor topic, we have adhered to the following steps,

a) For any particular topic, we have calculated sentiment score for each tweet belonging to that topic.
b) We then 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.
d) The value of |Neu| has been obtained using the formula: (1- (|Pos| + |Neg|)) (*as polarity value lies in [-1, 1]* (Section 4.3.1)).

Note, the method for evaluation of polarity score for each tweet has already been discussed in Section 4.3.1. Now, the controversial score is calculated using Equation 1. Table 6 depicts the controversial scores for rumor and non-rumor topics. It can be observed that, unlike non-rumors, the controversial score of rumor topics are much higher. This is due to the fact that, rumor events often trigger discussions, sentiments and opinions among the public. Such sort of outburst of people about the event often makes it controversial. However, in case of non-rumors, as the facts are acceptable for the people, they provide impersonal views about the events that make non-rumor topics less controversial.

Table 6: 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

5 Conclusion & Future Work

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 three rumor propagation features with proper illustrations. From the effects of each feature over rumors, we can effectively infer the correlation among these rumor propagation aspects. That is, *rumor events always warrant a discussion among the twitter community.* For such reason, the rumor topic remains alive through periodic discussions (*temporal*) where distinct user handles (*structural*) express their opinions / views through anger, anxiety, sorrow etc (*linguistic*) about the topic. 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 remained a burning issue. In our future study, we shall to concentrate on such issues through analyzing the effects of rumor propagation features on rumors and non-rumors for different disaster domains across the globe. Furthermore, we want to analyze which rumor propagation feature(s) is / are most suitable to detect rumors at an early stage in post disaster situations.

REFERENCES

[1] Pendleton, Susan Coppess. "Rumor research revisited and expanded." Language & Communication (1998).

[2] Oh, Onook, Kyounghee Hazel Kwon, and H. Raghav Rao. "An Exploration of Social Media in Extreme Events: Rumor Theory and Twitter during the Haiti Earthquake 2010." In Icis, vol. 231. 2010.

[3] Oh, Onook, Manish Agrawal, and H. Raghav Rao. "Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises." Mis Quarterly 37, no. 2 (2013).

[4] 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.

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

[6] 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.

[7] 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.

[8] 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, 2013.

[9] 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.

[10] Ma, Jing, Wei Gao, Zhongyu Wei, Yueming Lu, and Kam-Fai Wong. "Detect rumors using time series of social context information on microblogging websites." In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, pp. 1751-1754. ACM, 2015.

[11] 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.

[12] Hartigan, John A., and Manchek A. Wong. "Algorithm AS 136: A k-means clustering algorithm." Journal of the Royal Statistical Society. Series C (Applied Statistics) 28, no. 1 (1979): 100-108.

[13] Myers, Leann, and Maria J. Sirois. "Spearman correlation coefficients, differences between." Encyclopedia of statistical sciences 12 (2004).

[14] https://www.nltk.org/ (Last accessed on 07/10/18)