Isolating Rumors Using Sentiment Analysis

V. Sivasangari[1,*], Ashok Kumar Mohan[1], K. Suthendran[2] and M. Sethumadhavan[1]

[1]*TIFAC-CORE in Cyber Security Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India*
[2]*Kalasalingam Academy of Research and Education, Krishnankoil 626 126, Tamilnadu, India*
E-mail: cb.en.p2cys16022@cb.students.amrita.edu; {m_ashokkumar, m_sethu}@cb.amrita.edu; k.suthendran@klu.ac.in
**Corresponding Author*

Received 05 January 2018; Accepted 31 March 2018;
Publication 12 June 2018

Abstract

In recent days, social media has become a platform to spread false facts all the way through internet. One of the growing data analytic engine from web informal organization says, twitter has become the prime source for spreading fake news facilitating numerous perpetrators around the globe. It has turned into a competent, speedy cum effortless hotspot for news-fans to just click and forward junk data. Individuals are opting to use twitter for searching information regarding crisis circumstances and everyday occasions. In twitter, the spread of fraudulent or inaccurate information during the emergency situations will affect the individuals and public in numerous ways, but the original news is more reliable when it is declared by the news channels. So, the importance of proposed framework is categorized in three steps; Initially, Twitter Scraper is applied to scrape the vast volume of tweets and metadata from the collected set of tweets for the study on former Chief Minister (state of Tamil Nadu) death case controversy during 2016. Then the threshold value based on negative polarity of common tweets for the scraped data is calculated, once the tweeted texts are different from the threshold condition it will be automatically tagged as 'rumor' (negative) or else tagged as 'non-rumor'

Journal of Cyber Security and Mobility, Vol. 7_1, 181–200. *River Publishers*
doi: 10.13052/jcsm2245-1439.7113

(positive) using sentiment classifier. Finally, the proposed model on VADER based sentiment analysis identifies the false facts. It is obtained as a result of the sample tweets regularly training and testing it on whole datasets.

Keywords: Social Media Analysis, Twitter Scraper, Sentiment Classifier, Rumor, Non-rumor, *#RIP*.

1 Introduction

Majority of the Internet population in today's world have been addicted to social media networks such as Facebook, Twitter, Instagram and YouTube. Comparatively, the internet users spend more time on social networks than the search engines. People started building social media relationships [17] to improve direct communication with other online users. This kind of technology growth can be used to deduce each individual's social relationship with others through online social media networks. It solely depend on reaping all the end users information shared publicly via tweets. Even if the small catchy information is been disclosed online, then it will spread virally through social media networks without even knowing what exactly the information means. Thus the collected information that spreads across the social network is to be analyzed quickly and effectively to sort out the issues raised on the integrity of the data. This type of spreading unwanted propaganda is defined as an infection which has been spread among all the people in the vicinity of the tweet. However, at the same time social media networks become susceptible to different types of unwanted and malicious spammer or hacker actions merely committed for fun and profit. There arise a crucial need for the society and industry to deliver an appropriate solution in social media to restrict or report these false tweets.

Recently, most of the people have started believing the information shared in social media as authentic without even validating it once from any trusted source. Individuals have a right to know whether the data they are receiving is reliable or not. At the initial stage, when the malicious information starts spreading and it is difficult to predict the motivation of the trespasser. In most of the cases, the information that are shared at the time of emergency situations will be an inaccurate information and it will have enormous amount of negative impacts.

Twitter is one of the popular social media platform facilitating micro-blogging and has more than 330 million active users per month (as on September 2017). There will be a sudden increase in the user activity in twitter

during trending events like natural disaster, political events and bomb blast. Individuals are interested to check and update about the trending real world events by logging onto twitter or other social media websites. Since it acts as an open platform to the individuals to share information and also people can share their opinions about any trending events. Comparatively, twitter is providing huge, wealthy information over Facebook with lots of privacy restrictions. This kind of huge collection of information may have disproportionate mixture of true and false information. Different types of false contents like rumors, and gossips are considered to be a type of unwanted content which may make an issue at the time of disaster [13]. So before sharing the information in social media it is recommended to be aware of the source and nature of a the malicious information. Initially, from the huge collection of data, the original information or authentic and fraudulent information has to be separated to overcome this problem.

The main objective of the paper is to detect the rumor during twitter trending events when the tweets are tweeted and retweeted in a hurry without appropriate verification. One such immoral incident happened recently in the state of Tamil Nadu (India) is the former Chief Minister's death controversy. Based upon this case *#rip #cm* and similar tweets are sorted out. It is sequentially collected and recorded using Twitter Scraper. Datasets are collected from the day when the casualty has been admitted in a renowned hospital to the final day of the official announcement of her death confirmation (September 22, 2017–October 14, 2017). By considering the above mentioned case, the detection of rumor and malicious user who were spreading false news is being analyzed and classified.

The rest of the paper is organized as follows: The second phase is focused on the review about the problem of detecting rumors in twitter. The third phase explains about the developed framework over the case study with the proposed system and results. And finally in the fourth phase, the conclusion of the findings and future work of the paper was discussed.

2 Literature Survey

The semantic and sentiment analysis classifier was developed to find the false news on twitter. This classifier collects the twitter text, which can be used to verify the information or tweets from a standard account news channels during any disastrous event occurring in Twitter, this method is very much useful to find the rumors in the critical situation [1]. The training set of twitter content is also collected from the Boston bombing blast event using

Twitter API. Twitter API is better when compare to the other social media like Facebook while collecting the social media data for effective dataset creation. In Facebook API, it has more restrictions when compared with Twitter API while collecting the data. So, the number of Facebook datas such as shares, posts, likes, comments that can be collected [25] by the Facebook API user is minimal. Since this API has few restrictions while sharing the information via API, the spreading of fake news through the API is comparatively low. The importance of the examining the unwanted (fake) account and spreading of a false information is been examined appropriately [2]. The regression prediction model can be used to monitor and analyze the user activity such as isolating the individual who shared the unwanted tweets linked to *#boston-marathon* during the Boston marathon bomb blast to measure the growth of the fake information. Here they conclude with the result as 29% of tweets to be a false information, 51% of tweets marked as true and then remaining to be a set of common unclassified tweets. They have analyzed over the collected data and the outcome is demonstrated using temporal analyze technique that most of the viral texts are shared via mobile phones. The automatic detection and verification of rumors in twitter to overcome the social media affirmation is discussed in this work [3], which is classified based on a speech-act classifier. This classifier uses both semantic and syntactic features to detect the affirmation in social media with high precision. They eventually collect the conversations (tweets) which include a set of replies [7] during an emergency situation. Outcome result shows that, 209 tweets were fake over 9,38,806 tweet collected and this information is balanced with 75% accuracy. In order to demonstrate how many times that particular hashtag(#) has been shared repeatedly, the temporal analysis algorithm [2] is been implemented at this juncture [4]. Here they have made a graph to represents the iPhone related keyword and to examine the temporal distribution, plotted for time along X-axis and the actual count of tweets over Y-axis. Finally, they conclude that most of the retweeted text is outsourced via the suspect's iPhone based upon the graph results. Rumor can be defined as unpleasant tweets [3] or from an unverified source [6] and it is shared between one individual to the other via Twitter [5].

3 Proposed Framework

In online platform, e-crimes are increasing rapidly at the end of every hour. Majority of the individuals started believing the tweets without validating the information shared on social media. Due to these kind of issues, the

Figure 1 First original tweet officially announced in twitter about her death.

social media has gathered the attention of the data analysts to investigate and classify such malicious tweets. Here we consider the scenario of the former CM from the state of Tamil Nadu admitted at a hospital in the city of Chennai reported with some minor health issues on September 22, 2016 and she was declared dead on Dec 6, 2016. Starting from the month of September when she was hospitalized and till the date of her reported death, people were sharing enormous amount of tweets about her death every now and then. There are number of different malicious tweet sources who have shared fake news and fake tweets for more than six months. The following Figure 1. shows the interesting tweet isolated by our approach:

> Even before the hospital authorities released a confirmation about her death, rumors have started spreading before two days that she was unofficially declared dead by doctors. Actually the official news was announced by them only on December 6, 2016 exactly at 12:15 AM, or that is to say just forty five minutes after her reported death as shown in the Figure 1.
>
> There are a variety of tweets that has been shared on Twitter for this case and all have been collected precisely. So the paramount contribution of the paper is entirely focused on the differentiation of the malicious information (rumor), original news (non-rumor) and also about the impact of such unwanted malicious information in near future. The main contribution of the proposed work is listed in the Table 1.

3.1 Data Collection

The Datasets are collected with the reference of PHEME standard dataset [13]. The datas are collected from the day when the casualty has been admitted in

Table 1 Isolating rumors using sentiment analysis

Isolation Mechanism	Tasks to be Performed	Frameworks Used
Scraping the tweets with metadata	To scrape the appropriate # in Twitter	Twitter scraper python package 0.2.7
Identification of rumor using sentiment analysis method	i) To visualize the true tweets and malicious text	i) Sentiment viz
	ii) To differentiate the rumor and non-rumor	ii) Sentiment classifier python packages.
Demonstrate the fake retweets using using Temporal Analysis	To demonstrate fake retweet information and examine how malicious texts spreads rapidly	Plotly based Python API

a renowned hospital to the final day of the official announcement of her death conformation (September 22, 2017–October 14, 2017). PHEME dataset is made for finding the rumor in five different trending breaking news. The people seeing a course of events of tweets about the breaking news, a client would then clarify every one of the tweets similar to rumor or a non-rumor information. The pheme dataset is to assemble a dataset of bits of gossip (rumor) also, non-bits of gossip (non-rumor) was to build up a way to gather a different arrangement of stories, which would not really be known from the earlier and which would incorporate the two gossipy tidbits (rumor) what's more, non-gossipy tidbits (non-rumor).

In PHEME dataset, the five different real time incident namely Sydney Siege, Ottawa Shooting, Germanwings Crash, Charlie Hebdo, and Ferguson data (tweets) was collected with the help of Twitter API. The scraped tweets are used to find the information that are spreading was original or fake. The total number of tweets extracted in this dataset are 5802, in which 1,972 tweets are marked as rumor and 3,830 tweets are marked as non-rumor. In the Sydney siege incident, it is been found that 42% of information are rumor, Charlie Hebdo 22% of information are rumor and in Ferguson 24% of information are rumor. In Ottawa shooting they annotated 52% rumor and finally in the German wings Crash 50% rumor was found.

In most of the standard research works related to rumors, datasets (PHEME) were collected with the help of Twitter Streaming API [20] [22]. In order to provide authentication to the users, Twitter API is providing a (OAuth)[2] [3] access token to the user, so the user can read and access whatever the information is shared on Twitter in an authentic manner. Twitter Streaming API accumulates [19] the tweets for a particular circulating story with the main keywords based on particular hashtags(#). This API works based

on many standard protocols like OAuth combined to develop an custom-made application. Utilizing the streaming API, a persistent live stream session will be established between the server and the customer. It is mainly used to sniff and record all the data when it was publicly accessible and also to identify whether any new tweet are added to the user space. Twitter API has some constraint feature as the user will be limited to sent only 180 request for every fifteen minutes [26]. So the user can extract a maximum of 100 tweets per request and also it is limited for the user to access only the past seven days shared tweets. While Twitter does not give an API endpoint to recover discussions incited by tweets, it is conceivable to gather them by scratching tweets through the web customer interface. And also, for getting the accurate dataset we need a previous older data. But the pheme datasets are used a Twitter API, by using this API tweet crawlers cant able to scrape the previous years twitter data. But for the proposed investigation scenario, essentially we have to collect the older information from any social media [17]. To avoid the above limitation, tweets are collected and dataset is generated using the twitter scraper framework.

Twitter scraper provides the features to the user to read and access the past tweets. Twitter scraper will scrap all the user information such as username, location, twitter contents, screen name and time slots from Twitter. Here in this work, tweets with metadata is collected for the reported death case, which is shown in Figure 2. and scraped contents are stored in JSON output file. And also our datasets fed into the classification algorithm to check the accuracy of rumor detection. So, finally we got the 90% accuracy of our formal Tamilnadu CM Jayalalitha death case dataset.

Datasets are collected based upon popular hash tag that is been shared in Twitter related to the case. The sample hash tags are: *#Jayalalithaa, #ripamma, #ripjayalalithaa, #ammadeath, #purachithalaivi, #jayalal-itdeath, #ammaforever and #RIP.* These Tweets are collected and a precise dataset is formed to identify the rumor based on the proposed algorithm as shown below:

Algorithm 1: Textwise Tweets (CommonTweetsSet)
1. *for each* $\in$ *CommonTweetsSet do*
2. *CommonTweetsSet* $\leftarrow$ *collectTweets(Hashtag)*
3. *if carry obscure hashtag then*
4. *TextSet* $\leftarrow$ *findTweet(CommonSet)*
5. *for each Text* $\in$ *TextSet do*
6. *TextwiseCommonTweetSet* $\leftarrow$ *collect Tweets(text)*

7. *end for*
8. *end if*
9. *end for*
10. *return textwiseTweets Set*

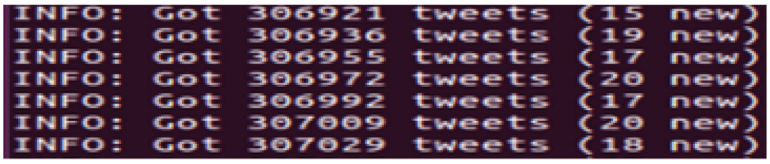

Figure 2 Download tweets using Twitter Scraper.

Based on the above hashtag, datasets are collected from Sep. 22, 2016 to Oct. 14, 2017 using Twitter scraper. To detect the rumor, the collected tweets are given as an input to the custom made tweeter scraper algorithm and specifically trained for this case.

3.2 Identification of Rumor

Social rumor can be defined as unauthentic accounts or explanations of events that circulates among the individuals on Twitter [9]. Most of the people believe in the fake information at first, but later they sense that that unwanted fake news is different from verified original news [4]. The original news channels are verifying the information before declaring the news to the outside world. To avoid a malicious source account, the news channel's authentic tweets are verified by the Twitter and the information or news from the original news channel account are considered as a trusted benchmark information [3, 12]. So the main objective of the following rumor detection algorithm is to differentiate the original news (non-rumor) [10] from the unwanted (rumor/fake) news [16].

Sentiment Viz tool is used to differentiate fake and true news based on the Natural Language Processing (NLP) techniques [18]. So once if the hashtag input is given by the user, it will separate the text based upon positive and negative tags and extracts the key text like happy, sad, unpleasant and relaxed. For example, here the hash tag (#rip) is given as an input, then it shows the

Figure 3 Sentiment analysis based rumors classified for the hashtags(#) using Sentiment Viz.

sentiment diagram for that corresponding hashtags. In the above Figure 3. the dots on the left side represents the 'fake messages' and the dots on the right side represents the 'true tweets' as shown in Figure 3.

The collected tweets are divided into two categories namely, the original news sets and common news sets. The verified tweets from news accounts tool are considered as original news while the remaining unnecessary tweets are fall into common tweet set. Based on the sentiment analysis algorithm [14], the news sets are classified as positive (non-rumor) or negative (rumor) and in some cases as neutral (unclassified). Then the sentiment polarity is also calculated for the above news sets based on the following formula:

$$News\ sets\ polarity = \begin{cases} positive\ \ if\ P\ >\ N \\ Negative\ \ otherwise \end{cases}$$

P = Total number of original news channel tweets with positive polarity
N = Total number of original news channel tweets with negative polarity

$$Threshold = Negative\ polarity/common\ tweets$$

$$Threshold = \frac{negative\_polarity}{common\_tweets}$$

Based on the polarity formula, if the P value is greater than N then it is considered as an positive polarity or else considered as an negative polarity.

Then, the verified tweets from the news channels and common tweets set from the end users are compared and analyzed, if the input tweets are

appropriate to the above threshold value then it is labeled as 'rumor' or else it ought to be labeled as 'non-rumor'[8]. Algorithm for finding the rumor is given below:

Algorithm 2: Predict (sentence)

Initialize Neg=0,Pos=0 ∈ sentence
for all words ∈ sentence do
  if word in word then
SR=sentiment.classifier (classifyWord)
  if SR= =Neg then
Neg ← neg +1
if SR= =Pos then
Pos ←pos+1
    end if
  end if
  end if
end for
for P=Pos mod (word) do
for N=Neg mod (word) do
if(P > threshold) then
  return 1
  else
  return 0
    end return
  end else
  end return
end if

From the above proposed algorithm 2 utilizes the sentiment classifier to categorize a tweet to be true (non-rumor) or false (rumor) based on the threshold value and tweet source.

Based on Table 2, the total number of tweets and retweets are classified into a rumor or non-rumor. If a text is identified as rumor (or fake news), it implies that the end users will be able to find the difference between original and unwanted news [10], impact during the important event [13] or a typical doubtful scenario as taken in our case. So this will help to reduce the occurrence of any such rumors at much earlier stages [11] of retweets.

Table 2 Top five results of sentiment analysis to scale rumor and non rumor

Selected Set of Hashtags (#)	Tweets/Retweets	Rumor	Non-Rumor
purachithalalivi	41,083	11,867	27,876
Rip jayalalitha	13,063	3960	9103
jayalaitha	3,08,470	2,51,238	57,232
amma death	3960	198	3762
Jayalalitha death	76,653	52,184	24,469

3.2.1 Analyze the Fake Retweet content

With the help of the temporal analysis, the retweet fake text is been analyzed. To examine the temporal distribution [2], the above listed counts in the Table 2 is been used for mapping the event. The below Figure 4. shows the growth of fake content based on retweets. From the above plot, it is been concluded that one of the hashtag result "Jayalalithaa killed" is rumor, based on the statistics from Table 2 and this hashtag tweets have been shared and retweeted many times compared to other hash tag in Twitter. Here it is also considered that if the retweet fake contents are retweeted many times then there will be a greater probability to affect the future occurrence of similar incidents. To analyze the fake retweet information, in the above Figure 4. the X-axis is marked as a source and Y-axis is denoted as a text. Initially, the tweets moves gradually and then after some time when the particular fake hashtag is getting retweeted by some malicious user, the tweets growth has increased progressively. Finally, once that particular retweeted fake information has become shared many times it spreads virally.

Identification of the source (tweet) of the rumor is a separate area of interest [15, 21] and the same has been justified using source estimators like Jordan centre to trace back the first individual(s) who started the rumor tweet at first [22]. It can also be mapped into a series of tweets in timeline to trace the rumor source.

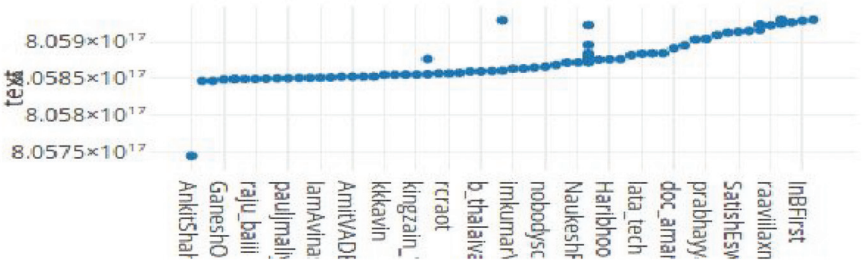

Figure 4 Analyze the retweet fake content growth.

4 Sentiment Analysis vs VADER Sentiment Analysis

To improve the accuracy, proposed a new approach called VADER sentiment analysis for isolating the rumor. Difference between sentiment analysis and VADER sentiment analysis are shown in Figure 5.

VADER sentiment analysis is a lexicon and rule-based sentiment analysis technique that is specifically attuned to sentiments expressed in social media, and works well on texts from other domains. Lexical method looks at the sentiment category or score of each and every word in the sentence (tweets text) and decides the category of the each word (with strength of the tweet) and also the nature of the text. The approach proposed for isolating the rumor can be classified into 3 phase which was shown in Figure 6.

4.1 Crawl Tweets with Metadata

The tweets are scraped from the twitter with the help of the Twitter Scraper and stored in the Json format. Json structure is maintained as a array structure which includes id, name, timestamp, text etc., The additional information of the text is extracted from the Json structure for the text based sentiment analysis method.

4.2 VADER Sentiment Analysis

VADER is one of the technique for text based sentiment analysis. It is classified into polarity (lexicon or dictionary) and rule (valence) based technique [24]. Polarity (lexicon) approach is used to categorize the positive, negative and

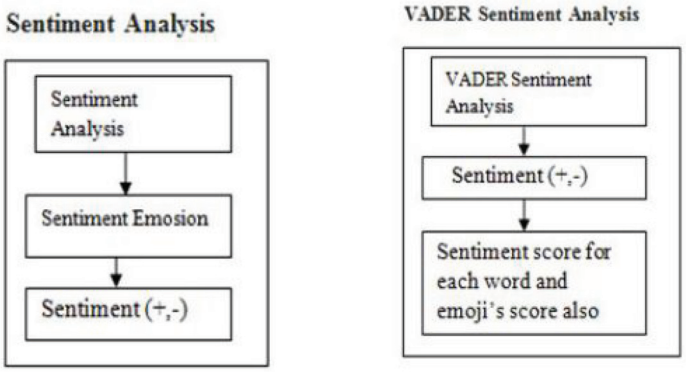

Figure 5 Sentiment Analysis vs VADER Sentiment Analysis.

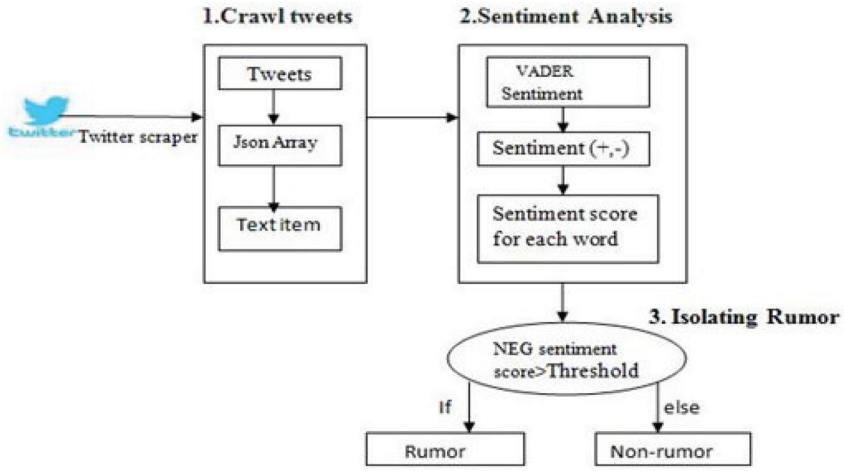

Figure 6 Architecture diagram for Isolating Rumor.

neutral. VADER is not only considering the sentiment category, its also considering the intensity (strength) [23] of the text using the rule based heuristics method with the value for each and every word in the text (tweets).

Pseudo code for Isolating Rumor using VADER:

```
#def sentiment_scores (self, sentiments)
pos_sum = 0
neg_sum = 0
neu_count = 0
for sentiment score in sentiments:
    if sentiment_score > 0:
        pos_sum += ((sentiment_score) + 1)
    if sentiment_score < 0:
        neg_sum += ((sentiment_score) - 1)
    if sentiment_score == 0:
        neu_count+=1
return pos_sum, neg_sum, neu_count
        End if
      End if
   End if
End for
THRESHOLD = 0.5
#analyzer = SentimentIntensityAnalyzer
```

```
all_text_score = []
for tweet in json_data:
    if 'text' in tweet.keys():
        text = tweet['text']
        tweet_count+=1
    if text:
        vs = analyzer.polarity_scores(text)
    if vs['neg'] >= NEG_THRESHOLD and vs['compound'] < 0:
        return (tweet_count, vs, rumor)
    else
        return (tweet_count, vs, non rumor)
            End for
           End if
         End if
       End if
End else
```

Based on the above Table 3, the extracted crawl texts (tweets) are isolated into rumor and non-rumor using VADER Sentiment Technique. Using VADER sentiment polarity and valence based technique, the extracted text are categorized as a positive and negative value for each crawl twitter text (tweets). Then, if the negative value is greater than the threshold (0.5) it will be marked as a rumor. If a text (tweets and retweets) is marked as a rumor, then the normal twitter user can able to differentiate the true information and

Table 3 Result for Segregating rumor

Tweets	Compound Value	Negative Score	Neutral Score	Positive Score	Rumor/ Non-Rumor
#sasikala MURDERED jayalalitha	–0.7297	0.72	0.28	0.0	Rumor
#Jayalalitha has been dead#IRONLADY	–0.7003	0.489	0.291	0.22	Non-rumor
#killed jayalalitha#AIADMK crisis	–0.6705	0.692	0.308	0.0	Rumor
#shocking sasikala killed jayalalitha–sasikala murdered	–0.9118	0.699	0.301	0.0	Rumor
#amma the strong women died because of cardiac arrest	–0.9223	0.403	0.471	0.126	Non-rumor
#shocking!!!JAYALALITHA HOST in HOSPITAL	–0.8943	0.595	0.405	0.0	Rumor
#J Jayalalithas death: 280 people have died grief, shock.	–0.9231	0.727	0.273	0.0	Rumor

rumor during the important incident. By consolidating these heuristics into VADER sentiment model, we radically refine the truth based on the accuracy of tweets.

5 Conclusion and Future work

Based on the above datasets collected, the Twitter Scraper has more advantage compared to the native Twitter API for this specific case study on rumor classification for a reported death. With the help of the Twitter Scraper, the anticipated scheme have scraped massive amount of tweets with metadata for this scenario. Sentiment analyze is used to differentiate the true and fake text for that large volume of scraped information and also mentions about how the retweeted fake text might affect any such similar occurrences in near future. And finally, found a sentiment lexicon score value using VADER for the scraped dataset to segregate the rumor with greater accuracy.

For the future work, a complete study of all social media networks and also an efficient methodology will be proposed for predicting the rumors in social media networks. We are working towards building a stable interface to investigate the datasets created on twitter to identify a fake identity. Also to sort out malicious user profiles is their initial stages before they gain the momentum to spread a rumor. This approach addresses the problem of predicting the first occurrence of a rumor and preventing them from spreading online. Minimizing the spread of unwanted misinformation by using appropriate methods to find and prevent the rumors will be useful for the one who wants to keep residents well informed on the tweets. Later, additionally it might be helpful for the law enforcement agencies to improve or standardize the security level at the time of deciding the appropriate actions to be taken to avoid or suppress a typical social media rumor like RIP.

References

[1] Jain, S., Sharma, V., and Kaushal, R. (2016). Towards automated real-time detection of misinformation on Twitter. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI),* (pp. 2015–2020). IEEE.

[2] Gupta, A., Lamba, H., and Kumaraguru, P. (2013). $1.00 per rt# boston-marathon# prayforboston: Analyzing fake content on twitter. In *eCrime Researchers Summit (eCRS)*, 2013 (pp. 1–12). IEEE.

[3] Vosoughi, S. (2015). Automatic detection and verification of rumors on Twitter (Doctoral dissertation, Massachusetts Institute of Technology).
[4] Xia, F., Yu, C., Xu, L., Qian, W., and Zhou, A. (2017). Top-k temporal keyword search over social media data. *World Wide Web,* 20(5), 1049–1069.
[5] Shah, D., and Zaman, T. (2011). Rumors in a network: Who's the culprit?. *IEEE Transactions on information theory,* 57(8), 5163–5181.
[6] Liu, S., and Young, S. D. (2016). A survey of social media data analysis for physical activity surveillance. *Journal of Forensic and Legal Medicine*.
[7] Zubiaga, A., Liakata, M., Procter, R., Bontcheva, K., and Tolmie, P. (2015). Towards Detecting Rumours in Social Media. In *AAAI Workshop: AI for Cities*.
[8] Zhang, D. Y., Han, R., Wang, D., and Huang, C. (2016). On robust truth discovery in sparse social media sensing. In *Big Data (Big Data), 2016 IEEE International Conference on* (pp. 1076–1081). IEEE.
[9] Zheltukhina, M. R., Slyshkin, G. G., Ponomarenko, E. B., Busygina, M. V., and Omelchenko, A. V. (2016). Role of Media Rumors in the Modern Society. *International Journal of Environmental and Science Education,* 11(17), 10581–10589.
[10] Wang, S., and Terano, T. (2015). Detecting rumor patterns in streaming social media. In *Big Data (Big Data), 2015 IEEE International Conference on* (pp. 2709–2715). IEEE.
[11] Mitra, T., Wright, G. P., and Gilbert, E. (2017). A parsimonious language model of social media credibility across disparate events. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 126–145). ACM.
[12] Arif, A., Robinson, J. J., Stanek, S. A., Fichet, E. S., Townsend, P., Worku, Z., and Starbird, K. (2017). A Closer Look at the Self-Correcting Crowd: Examining Corrections in Online Rumors. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 155–168). ACM.
[13] Zubiaga, A., Liakata, M., and Procter, R. (2016). Learning Reporting Dynamics during Breaking News for Rumour Detection in Social Media. *arXiv preprint arXiv:1610.07363.*
[14] Vinodhini, G., and Chandrasekaran, R. M. (2012). Sentiment analysis and opinion mining: a survey. *International Journal,* 2(6), 282–292.
[15] Louni, A., Santhanakrishnan, A., and Subbalakshmi, K. P. (2015). Identification of source of rumors in social networks with incomplete information. *arXiv preprint arXiv:1509.00557*.

[16] Pasquini, C., Brunetta, C., Vinci, A. F., Conotter, V., and Boato, G. (2015). Towards the verification of image integrity in online news. In *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW),* (pp. 1–6). IEEE.

[17] Mohan, A. K., and Venkataraman, D. (2017). Forensic future of social media analysis using web ontology. In *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS),* (pp. 1–6). IEEE.

[18] Sanjay, S. P., Anand Kumar M, and Soman, K. P. (2015). AMRITA_CEN-NLP@ FIRE 2015: CRF Based Named Entity Extractor For Twitter Microposts. In *FIRE Workshops* (pp. 96–99).

[19] Wang, M., and Gerber, M. S. (2015). Using Twitter for Next-Place Prediction, with an Application to Crime Prediction. In *Computational Intelligence, 2015 IEEE Symposium Series on* (pp. 941–948). IEEE.

[20] Zhao, Z., Resnick, P., and Mei, Q. (2015). 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.

[21] Luo, W., Tay, W. P., Leng, M., and Guevara, M. K. (2015). On the universality of the Jordan center for estimating the rumor source in a social network. In *2015 IEEE International Conference on Digital Signal Processing (DSP),* (pp. 760–764). IEEE.

[22] Krithika, R., and Mohan, A. K. Inspecting Irresponsible Hypes: Rumors in Social Media Networks.

[23] Fazal Masud Kundi1, Lexicon-Based Sentiment Analysis in the Social Web.Institute of Engineering and Computer Sciences, Pakistan.

[24] Gilbert, C. H. E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International Conference on Weblogs and Social Media (ICWSM-14).* Available at (20/04/16) http://comp. social. gatech. edu/papers/icwsm14. vader. hutto. pdf.

[25] https://stackoverflow.com/questions/8713241/whats-the-facebooks-graph-api-call-limit

[26] https://stackoverflow.com/questions/1285666/twitter-api-limit

Biographies

V. Sivasangari is pursuing her M.Tech. in Cyber Security at TIFAC-CORE in Cyber Security, Amrita Vishwa Vidyapeetham, Coimbatore, Tamil Nadu, India. Her current area of research is Rumor Prediction in Social Media Networks.

Ashok Kumar Mohan, M.Tech. specialized in Cyber Security, is a Research Associate at TIFAC-CORE in Cyber Security, Amrita Vishwa Vidyapeetham, Coimbatore, Tamil Nadu, India. He is currently a Ph.D. scholar doing his research in the area of Cyber Forensics funded by Ministry of Electronics & Information Technology (Government of India) under Visvesvaraya PhD scheme for Electronics and IT. He is currently pursuing his research over the cyber security core vicinity in Metadata Forensics, Wireless Security Auditing, Rumor Prediction in Social Media Networks and Slack Space Analysis of NTFS File Systems. He is also the Certified EC-Council Instructor (CEI) for ethical hacking and penetration testing certification courses at the research centre.

Suthendran Kannan, received his B.E. Electronics and Communication Engineering from Madurai Kamaraj University in 2002; his M.E. Communication Systems from Anna University in 2006 and his Ph.D. Electronics and Communication Engineering from Kalasalingam University in 2015. He was a Research and Development Engineer at Matrix view Technologies Private Limited, Chennai for a couple of years. He is now the Head, Cyber Forensics Research Laboratory and Associate Professor in Information Technology, Kalasalingam Academy of Research and Education. His current research interests include Cyber Security, Communication System, Signal Processing, Image Processing, etc.

M. Sethumadhavan, received his Ph.D. (Number Theory) from Calicut Regional Engineering College. Currently, he is working as a Professor in the Department of Mathematics and Computer Science, Amrita Vishwa Vidyapeetham University, Coimbatore. His current research interests include: Post Quantum Crytography, Block Chain and Boolean functions.