

# Rumor Detection In Social Media

Sumesh Meppadath<sup>1</sup>, Kaustubh Kansara<sup>2</sup>, Nishant Shrivastav<sup>3</sup>, Karen Joy<sup>4</sup>

<sup>1,2,3,4</sup> Final Year Student, Computer Department, K.K.Wagh Institute of Engineering Education and Research, Nashik, Maharashtra, India

## ABSTRACT

*Social Media has outpaced the conventional news media systems. It is often observed that news stories are first broken in cyber space and then the electronic and print media take them up. However, the distributed architecture and lack of moderation in most of the social media platforms, with the temptation of the users for posting a newsworthy story early on, makes the veracity of information a major issue. This information is defined as rumor, which is a non-credible piece of data circulating in cyberspace, often causing social unrest. For establishing credibility of this information, we define a two phase approach, considering Twitter as the target social networking platform specifically of Indian domain. First phase is based on the premise that verified News Channel Handles in Twitter would furnish more credible information as compared to the nave general public at large. Livestreaming or recent tweets are extracted corresponding to Twitter trends, based on clustering using Hashtags. Contextual and sentiment mismatch ratio between the tweets of above mentioned classes is done using semantic and sentiment analysis of tweets, which would reflect the degree of discrepancy of the information. However, as a tweet is restricted to 140 characters, it may not be sufficient to gain useful insight. Tweets are also is susceptible to noise, which may decrease the accuracy of analysis. To overcome these problems, the second phase verifies the claim credibility result, from heterogeneous data resources like authentic web articles, web pages, blog posts, websites, etc. using web crawling techniques. The performance of the proposed method is evaluated, considering the recent examples of popular rumors that went viral in social media.*

**Keywords:** - Rumor Detection, Web Content Mining, Twitter

## 1. INTRODUCTION AND MOTIVATION

The motivation for doing this project was primarily an interest to undertake a challenging project in an area that helps a social cause. The public unrest that is caused when any misinformation spreads via social networks, is one of the rising issues which affects society's health and harmony. Implementation of the knowledge of computer science in solving an issue of society and the opportunity to develop a solution in collaboration with Nashik Cyber-security police to help the general public was the main motivation towards development of this project

### 1.1 Project Scope and Objectives

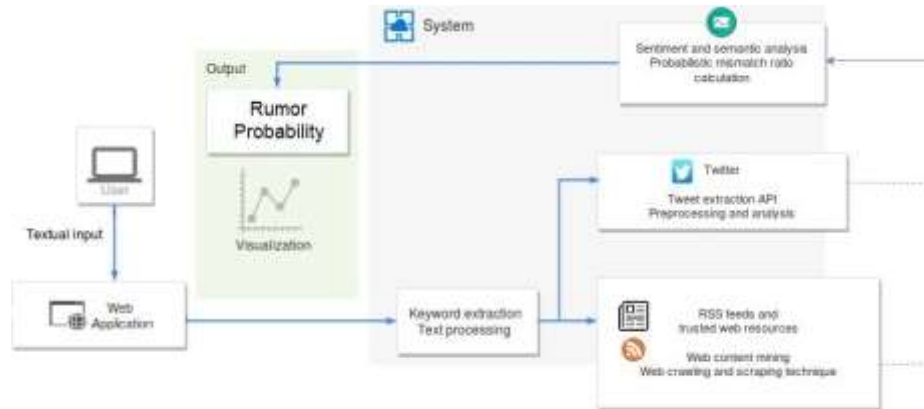
This system will classify information into rumor or fact. Keywords will be extracted from the provided textual content, which will be used to for further analysis.

Currentscopeislimitedonlytotextualcontentasaninputinbringingaprobabilistic conclusion to the input given by the user. In future, this project can be extended to accept user input in image or video format. Furthermore, it can be integrated into any social media platform for custom surveillance of content. The objectives are as follows:

- To provide a platform for general public to know a fact from rumor
- To help the Cyber Cell to mitigate the effect of rumors on society
- To provide a platform for verification of non-credible information

## 2. SYSTEM ARCHITECTURE

The system architecture is designed in a way to work on real time data from twitter and heterogeneous web content, based on the keywords extracted from the user query. This enables to scan and process any information, compare it to authentic sources and bring forth a probabilistic answer.



**Figure -1:** System Architecture

The various components and the major implementations as shown in Figure 1 are as follows:

- **Web Application:** User inputs query through this to know the probability of the rumor
- **Keyword Extraction:** Important keywords are extracted from the user query using keyword extraction algorithm
- **Web-Content Mining:** Collecting Public tweets from twitter and verified information from heterogeneous web resources
- **Semantic and Sentiment Analysis:** Semantic Analysis is used to determine contextual similarity between the tweets and verified information whereas Sentiment Analysis is used for determining the opinion of people
- **Calculating Rumor Probability:** Determining the probability of rumor using mis-match ratio

Along with this main Rumor probability check system, two additional systems have also been implemented as follows:

1. **Fact Check:** Checking if the User Query is a fact or not and providing Headline, URL and content from the heterogeneous sources.
2. **Hot 5 News:** Displaying Hot5 articles which are responsible for debunking misinformation that we encounter on a daily basis on social media as well as mainstream media.

### 2.1 Mathematical model:

Input:  $s_1$  and  $s_2$

Where  $s_1$  and  $s_2$  are sentences in the form of word vectors

Output:  $Fos = (d * Fcs) + ((1-d) * Fwo)$

$Fcs = \cos^{-1}(s_1 \cdot s_2 / (\sqrt{s_1} * \sqrt{s_2}))$

$Fwo = |j_1 - r_2j| / |j_1 + r_2j|$

$Fmismatch = 1$  if  $Fos \geq \text{Threshold}$   
 $= 0$  otherwise

$Frp = M/T$

Where  $Fos$  - Overall Similarity

$Fcs$  - Cosine Similarity

$Fos$  - Word Order Similarity

Frp - Rumour Probability

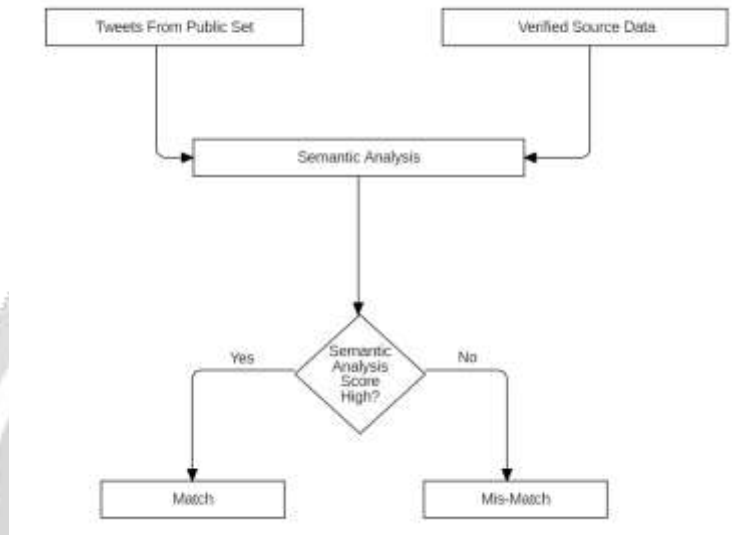
r1 and r2 are best tokenized word vectors

M - Number of tweets with polarity opposite to newsPolarity

T - Total number of public tweets

d - 0:62 0:35<Threshold<0:60

## 2.2 Rule engine:



## 2.3 Algorithm

Algorithm calculateSentenceSimilarity(sentence1, sentence2)

```

{
  1. Tokenize the two sentences sentence1 and sentence2 into ts1 and ts2
  2. Create joint word set of two tokenize sentences ts1 and ts2, the result is treated as joint_word_set
  3. Calculate cosine similarity between two tokenize sentences ts1 and ts2, the result is treated as Ss
     Ss = calculateCosineSimilarity(ts1, ts2)
  4. Calculate word order similarity which is represented as Sr
     Sr = calculateWordOrderSimilarity(ts1, ts2, joint_word_set)
  5. Determine overall similarity by using a CONST_DELTA that determine the contribution of each
     similarity to overall similarity
     Overall_Similarity(S) = CONST_DELTA*Ss + (1-CONST_DELTA)*Sr
  6. return Overall_Similarity(S)
}
  
```

Algorithm calculateCosineSimilarity(ts1,ts2)

```

{
  1. Create two semantic vectors sem_vec_one and sem_vec_two from ts1 and ts2
  2. Calculate cosine angle between sem_vec_one and sem_vec_two by doing operation:
     Ss = dot_product(sem_vec_one,sem_vec_two)/(sqrt(magnitude(sem_vec_one) +
     magnitude(sem_vec_two)))
  3. return Ss
}
  
```

}

Algorithm calculateWordOrderSimilarity(ts1, ts2, joint\_word\_set)

```
{
    1. Assign index number to each word vector from the joint word set
    2. Get the best words from ts1 and joint_word_set by checking if its greater than a preset
threshold(CONST_ETA) which gives r1
    3. Get the best words from ts2 and joint_word_set by checking if its greater than a preset
threshold(CONST_ETA) which gives r2
    4. Calculate word order similarity by doing operation:
         $Sr = |r1 - r2| / |r1 + r2|$ 
    5. return Sr
}
```

## 2.1 Applications:

The output results indicate the trending rumor event (with high probability as true and low as false) as early as possible when needed in real-time environments. This system provides an easy way for the common public to verify and validate news and information. Also, complete implementation of this system could result into increase in knowledge and current affairs enabling the data to be gathered, stored and used for further analysis in need of crisis or for investigations.

## 3. RESULTS

The results have been given below in section 3.1 and 3.2 explaining the rumors that were considered manually in this system in check with accuracy and finally as displayed on the front end.

### 3.1 Analysis for Accuracy

We implemented the proposed algorithm and were able to gather some rumors.

1. One of the rumors, we found, was Michael Browns killing was a criminal-act. Prosecutors would have faced the hurdle of proving that Darren Wilson deliberately committed a racially motivated hate crime. We manually checked it and it was a verified rumor. On August 9, 2010, Michael Brown, an 18-year-old African American man, was fatally shot by 28-year-old Ferguson police officer Darren Wilson in the city of Ferguson led to months of unrest and revived a debate on race and law enforcement in the US. A grand jury in St Louis decided not to indict Wilson on state charges as there is no evidence upon which prosecutors can rely to disprove Wilson's stated subjective belief that he feared for his safety. The probability of being a rumor by implemented method is 87%
2. Another such rumor was, in the 2014 Sydney hostage crisis, also known as the Sydney siege that Sydney radio host Ray Hadley reported that there were terror raids in Lakemba, a suburb with a large Muslim population. We checked it manually and it was a verified rumor. On December 15-16, 2014, A lone gunman Man Haron Monis, held hostage ten customers and eight employees of a Lindt chocolate cafe in the APA Building in Martin Place in Sydney, Australia. But the truth turned out to be more prosaic it was a pre-arranged police tour of a mosque. There were no raids. The probability of being a rumor by implemented method is 62%

**Table -1:** Accuracy Analysis for Michael Brown Rumor Topic [1]

		Predicted Label	
		Match	Mis-match
Actual Label	Match	18	116
	Mis-match	52	198

$$\text{Accuracy} = \frac{18+198}{284} = 76.05\%$$

The results are better in this set of tweets since tweets are more objective and less subjective. Thus, the number of tweets with both semantic similarity and objectivity increases, increasing the number of correct classification.

**Table -2:** Accuracy Analysis for Sydney Siege Rumor Topic [2]

		Predicted Label	
		Match	Mis-match
Actual Label	Match	58	23
	Mis-match	40	93

$$\text{Accuracy} = \frac{58+93}{219} = 70.56\%$$

The accuracy is low compared to the Mike Brown Rumor due to most of the tweets being subjective to and very less number of tweets showing objectively

#### 4. CONCLUSIONS

In conclusion, given access of the web portal to each and every one especially the general public will result in increasing awareness of real news versus the rumors that spread through social platforms. Even the authorities may take required action if they find an un-pleasant news to be a real time fact. Through our work, we implemented a working prototype which gives rumor probability using public tweets on Twitter and the verified news sources as base. Apart from rumors, general public can also check fact and HOT5 viral news which could result into increase in knowledge and current affairs. We believe that the system can be used especially in the critical times of emergency and can prove to be very useful to monitor and hence take actions to curb the spread of unvaried information.

#### 5. REFERENCES

- [1]. Suchita Jain, Vanya Sharma and Rishabh Kaushal, "Towards Automated Real Time Detection of Misinformation on Twitter", 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 2124, 2016, Jaipur, India
- [2]. Koichi Sato, Junbo Wang, Zixue Cheng, "Design of a Method to Support Twitter based Event Detection with Heterogeneous Data Resources", 2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST 2017)
- [3]. Shihan Wang, Takao Terano, Izabela Moise, Dirk Helbing, "Early Signals of Trending Rumor Event in Streaming Social Media", 2017 IEEE 41st Annual Computer Software and Applications Conference
- [4]. Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang and Huan Liu, "Fake News Detection on Social Media: A Data Mining Perspective", 2016 IEEE 41st Annual Computer Software and Applications Conference