



# Rumor Source Identification in Social Networks Using Time Varying Topology

Sherin Zibiah .S<sup>1</sup>, Sivaranjani.S<sup>2</sup>, Agnes Joshy.S<sup>3</sup>

<sup>1</sup>(Dept. of IT, UG Scholar, Francis Xavier Engineering College, Tirunelveli)

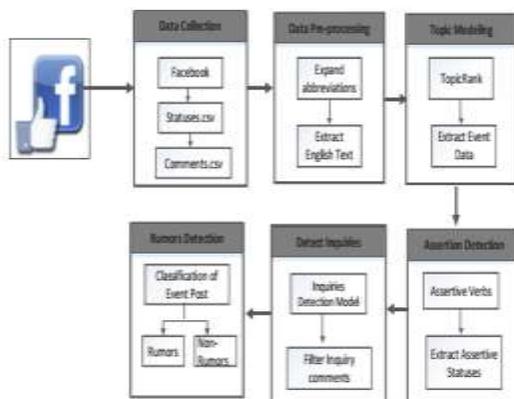
<sup>2</sup>(Dept. of IT, UG Scholar, Francis Xavier Engineering College, Tirunelveli)

<sup>3</sup>(Dept. of IT, Assistant Professor, Francis Xavier Engineering College, Tirunelveli)

**Abstract:** Identifying rumor sources in social networks plays a critical role in limiting the damage caused by them through the timely quarantine of the sources. However, the temporal variation in the topology of social networks and the ongoing dynamic processes challenge the traditional source identification techniques that are considered in static networks. The idea from criminology is borrowed and novel method is proposed to overcome the challenges. First, the time-varying networks is reduced to a series of static networks by introducing a time-integrating window. Second, instead of inspecting every individual in traditional techniques, we adopt a reverse dissemination strategy to specify a set of suspects of the real rumor source. This process addresses the scalability issue of source identification problems, and therefore dramatically promotes the efficiency of rumor source identification. Third, to determine the real source from the suspects, a novel microscopic rumor spreading model is developed to calculate the maximum likelihood (ML) for each suspect. The one who can provide the largest ML estimate is considered as the real source.

## I. INTRODUCTION

The discovery of social network services has led to the public to spread rumors at fastest rate. Thus, analyzing user's behavior on microblogging platforms to reach credibility of such information and examined the retweets system to analyze the rumors propagation pattern on the Twitter.



Qazvinian et al. classified the rumors related tweets using the matching regular expression with the keyword query. Zhao et al. approach based on the assumption that rumors will provoke tweets from the user's inquiring about their reliability. It implies that such tweets are possibly rumor having a number of enquiring tweets. He

prepared a list of five regular expressions that are useful to identify the inquiry tweets.

There has been found little work on automatic rumor detection regardless of the extensive study to examine the rumors in social media and developing techniques to tackle this problem. A set of predefined rumors is fed to a classifier, which classifies new tweets as being linked to the predefined rumors or not.

Public comments on the Facebook contain typos, misspelled words, unstructured and informal text. It is designed an approach to tackle the typos by expanding the abbreviations to correct the short form of words and recognize the English text using dictionary-based approach. It is created an abbreviation list of most commonly used abbreviations. It is compared each word in a comment with the abbreviation list to extract full form of word and replaced it in the actual dataset. To separate English text, a dictionary-based approach is applied in which each word in the comment is checked in the English dictionary and extracted the English text in a separate file. For each comment, number of English words are counted and divided by the total number of words present in the comment to get its weight as presented.

### Topic Modelling for key extraction

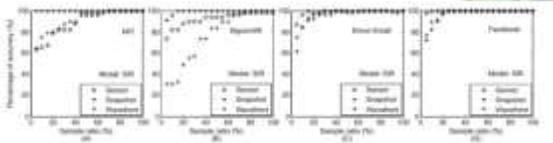
Topic Rank is an unsupervised, graph-based key phrase extraction method. In Social media, event detection is a prominent research topic. Topic Rank is used to discover



the topics that need to be manually analyzed during post processing to select the best topic, describing about the events.

It has become a challenge to obtain the event relevant posts, since posts may have event relevant terms but describing something other than the event such as a post containing the term “earthquake” could refer to an actual earthquake or to a conference on earthquakes.

Generally in a document, one noun phrase is sufficient to convey the topic. Therefore, some candidate keyphrases are redundant to represent the topic. Existing graph-based methods do not take that fact into account. Candidate weighting, is assigned using a random walk algorithm. N-best selection, keyphrases contains the 10 highest scored candidates as (keyphrase, score) are extracted.



Having an social resolver of rumors, It would have been liked to devise an algorithm that can predict the veracity of rumors before they are officially resolved by Wikipedia. It can operationalized our rumors using the features discussed earlier.

## II. SCREEN SHOTS



Start up page .In this page detection can be started.



Reducing the time-varying networks to a series of static networks.



Adopting an reverse dissemination strategy to specify a set of suspects of the real rumor source.



Determining the real source from the suspects





calculate the maximum likelihood (ML) for each suspect. The one who can provide the largest ML estimate is considered as the real source.

### III. CONCLUSION

The current work explored the problem of rumors detection based on inquiry comments identification using textual content of social media especially Facebook. Data of Facebook is secure and inaccessible to access except the public pages.

### REFERENCES

- [1]. Eytan Bakshy, Jake M Hofman, Winter A Mason, and Duncan J Watts. Everyone's an in uencer: quantifying in uence on twitter. In Proceedings of the fourth ACM international conference on Web search and data mining, pages 65{74. ACM, 2011.
- [2]. David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. the Journal of machine Learning research, 3:993{1022, 2003.
- [3]. Clint Burfoot and Timothy Baldwin. Automatic satire detection: Are you having a laugh? In Proceedings of the ACL-IJCNLP 2009 conference short papers, pages 161{164. Association for Computational Linguistics, 2009.
- [4]. Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. Information credibility on twitter. In Proceedings of the 20th international conference on World Wide Web, pages 675{684. ACM, 2011.
- [5]. The Pew Research Center. Internet overtakes newspapers as news outlet, De-cember 2008. <http://pewresearch.org/pubs/1066/internet-overtakes-newspapers-as-news-source> [pewresearch.org; posted 23-December-2008].
- [6]. The Pew Research Center. Public evaluations of the news media: 1985-2009. press accuracy rating hits two decade low, September 2009. <http://www.people-press.org/2009/09/13/press-accuracy-rating-hits-two-decade-low/>[people-press.org; posted 13-September-2009].