

Prioritization of Key Objectives During Floods

V. Kakulapati

Abstract: Now a day social networks generates large volume of data per sec and one of such network is Twitter. Twitter is one of the popular public platforms with an extract of openly express user's opinion. Our work aims focus on tweets generated in regard to floods and especially the tweets posed by those affected by floods so that we may prioritize objectives in order to facilitate aid and relief to those affected people. This task is accomplish by identifying the needs and requirements of the survivors of these calamities using responses via twitter analysis, these needs and requirements are certain objectives such as provisioning of food, tents for people, etc., all of these objectives can be prioritize based on certain words used by the survivors and transforming into tokens. These token are called as lexical normalization. In this work we analyze the lexical normalization of data generated by twitter by applying various techniques and visualize the investigations as the techniques are applied to process raw data from Twitter.

Index Terms: Priority, Lexical, Tweets, Floods, Token, Opinion.

I. INTRODUCTION

Human, animal lives, crop and properties are destroying by the influence of natural disasters all round the globe. In all these situations, the motives are not in human control. At the same time as the message in [1], past three decades such 1970-80 i.e., natural disaster 2nd rank, 1980-90 i.e., natural disaster 4thrank, 1990- 00 i.e., natural disaster 2ndrank, Bharath ranks in the top five countries intensively the total loss of humans. The impact of natural disaster not only the instantaneous consequence as examined in [2], a revelation to a natural calamity in the precedent augments the chances of severe infections for instance, diarrhea, dengue fever, malaria and keen respiratory infection in kids below 5 years age between 9–18%. The major natural disasters are Earthquake, Landslide, Cloudburst, Storm, the flood, and Volcanic Eruption. Floods are an excess of massive water accumulations beyond average limits in excess of dry ground. Millions of individual lives, cattle and rural crops are wiped out per year in India owing to be deficient in planning and unacceptable weather predicting.

Twitter is a social media network, it is multifaceted provision that facilitate its members to post and comprehend short 280-character communications called "Tweets" and it also allow its users to repost someone else's tweet is called "Retweets". Tweeter users can comprehend as well as posing tweets, but those who are not twitter users can only view them but not make any post of their own. As of January 2018, Twitter has more than 330 million monthly dynamic users and over 500 million postings in terms of tweets generated per day.

Manuscript published on 30 December 2018.

* Correspondence Author (s)

V. Kakulapati, Department of Information Technology, Sreenidhi Institute of Science and Technology (SNIST), Yamnampet, Hyderabad (Telangana) India-501301,

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

From people who were affected by the 2017 Texas floods and many other since also there has been new technology developed around We have chosen to work with twitter as it was heavily relied upon by the news media to cover and get in-depth information this concern that help provide internet connectivity during such calamities such as project loon [15,16]. The data generated in twitter is the relevant data is large volume of data for twitter, as evaluated to conventional micro blogging sites. In addition, twitter allows non-registered users to still be able to read tweets made by the person who helps and works in perfect cohesion with our work the response on twitter is also quicker in terms of collecting information.

II. RELATED WORK

India is exposed to diverse natural disasters appropriate to its unique climatic situations. Due to India is bound by water on its east, west and south sides and the north side is the Himalayas. Every year the country has been strike roughly by 8 natural disasters, and about 5 times rise in frequency of disasters since last three decades. These natural disasters affect India can be classified as 60%island is lying face down to earthquakes, 12% floods and 8% are lying face down to cyclones. India is a sufferer of more natural catastrophes every year and the loss of human survives and assets add together to lots of money which India cannot have the funds for loss. The following reasons are the poor natural disaster management process pursued by India.

- Early warning methods are insufficient
- When disaster happens there is no proper management to facing the disaster
- Inappropriate and time-consuming relief process
- Improper administration
- Time-consuming rehabilitation and reconstruction
- Finance managements are very poor for attending relief work.
- Poor help for disaster affective lives.

Although twitter message contents are more casual and partial [3], many investigations utilized twitter data for the natural disaster location [4], entity tracking [5], and incident identification [6].Natural disaster related [7] information communicated messages with the intention of recognize the point of catastrophes and investigated locations revealed in disaster associated messages revealed locations so as to identify the spot of natural calamities and influenced areas. Though, categorizing thing or incident associated communications accurately is a difficult assignment. A system [8] developed which tracks the faction of seismic activities and hurricanes based on user posts identified on Twitter. In this system comparing the features of the event associated messages for building classifier, with a more efficient feature set achieving precision of 64%.

Prioritization of Key Objectives During Floods

In other investigation [6] twitter data utilizes for detecting crime and disaster incidents. In their work, the twitter messages are classified words detected in disaster related messages contained in the twitter data and attained a precision and a recall are 30% and 85 % respectively.

Researchers are working on numerous methods by using Lexicon analysis.

In these methods evaluate polarity of opinions taking as a function which contains sentimental text in tweets.[9], [10].

Lexicon backed rule based classification [11] methods have used to classify user reviews, which integrates consequence of emoticons, modifiers, negations, etc to the lexicon based framework to get better performance.

III. FRAMEWORK

This work aims to prioritize objectives in order to help provide aid and relief to parties affected by floods.

3.1. Lexical Analysis

This process converts a text to a chain of tokens. Diverse applications, such as text editors, search engines, machine learning algorithms and compilers utilize Lexical Analyzer to retrieve significance tokens while eliminating unnecessary white spaces, comments, etc. [12]. We utilize this method in our work for sentiment analysis of posts from twitter.

This task is accomplished in 3 steps, step-1 acquiring the tweets from twitter step-2 analyzing these tweets based on tokens, these tokens are of further 3 types (type 1-that specify immediateness ex- need, must, imperative, etc. The type 2-that specify non-immediateness ex- want, like, etc., and finally type 3-which determine the objectives which are food, potable water & shelter) step-3 once the analysis is finished the objectives based on urgency are displayed in forms of graphs namely bar graph and pie chart as a result. The framework used to do this is R Studio for analysis and Shiny web application for data representation.

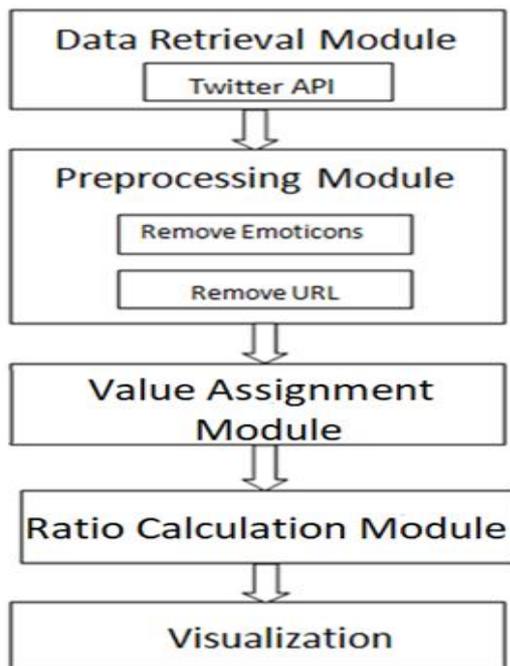


Fig 1. Framework of Proposed Method

Evaluating pre-defined tokens, after which retrieve tweets from twitter which is dynamic in nature based on the key

words provided on the server end, once these data sets are loaded, we begin to pre-process our data in order to remove any noise such as emoticons and URLs. Then execute the actual analysis invoking the value assignment module, in which based on tokens existing in each objective value is assigned based on each tweet mentioning the objective using an immediate or non-immediate word (tokens). The result is generated here this ends the value assignment segment (lexical analysis), the result mentioned earlier is recorded then further processed to generate a percentage this is the function of the Ratio calculation module which is used in plotting the pie chart, finally the visualization part of the code takes over and generates easily understandable graphs, one pie chart which uses the result stored as percentage values and a bar graph which uses the result stored as individual objective value so that the user can get a comprehensive understanding of the situation.

IV. IMPLEMENTATION RESULTS

Retrieving a data set from twitter which is dynamic in nature based on the key words provided on the server end once these data sets are loaded, we begin to remove any noise such as emoticons and URLs after this, based on tokens mentioned earlier each objective value is incremented based on each tweet mentioning the objective using an immediate or non-immediate word (or words) then the result is recorded, a ratio of individual objectives is taken with respect to others and is stored separately finally a graph for both results is plotted and visualized for the user.

The task of Tokenizer is conversion of the input text. Whenever new token come across is then counterpart for the lexicon in the vocabulary, which is found then the score is added to the entire pool of score for the contribution wording. Such as if “dramatic” is an affirmative match in the lexicon then the total text score is incremented or else the text score is reduced or the word is labeled as negative. However, this method emerges to be amateur in nature, its alternatives have demonstrated to be precious [13, 14]. In Fig. 1 illustrates the functioning of a lexical method.

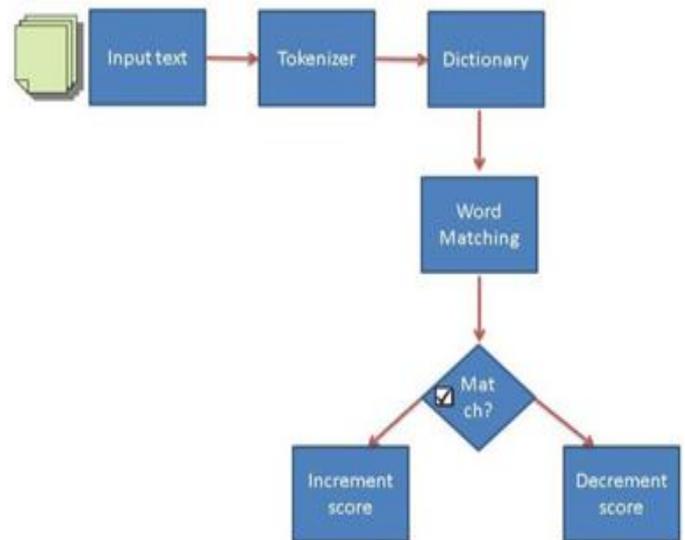


Fig 2. Working of a Lexical Technique.



Twitter Flood Analysis



Fig 3. Loading Tokens

Pie Chart

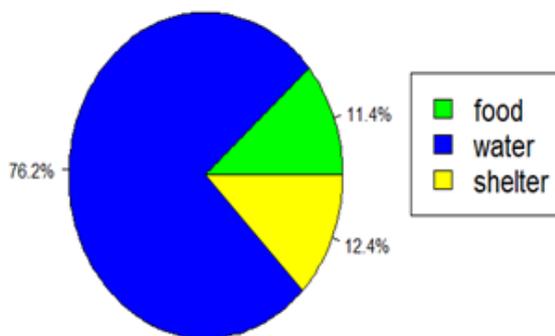


Fig 4. The Prioritization of Objects Which Contains in Twitter Message Tokens

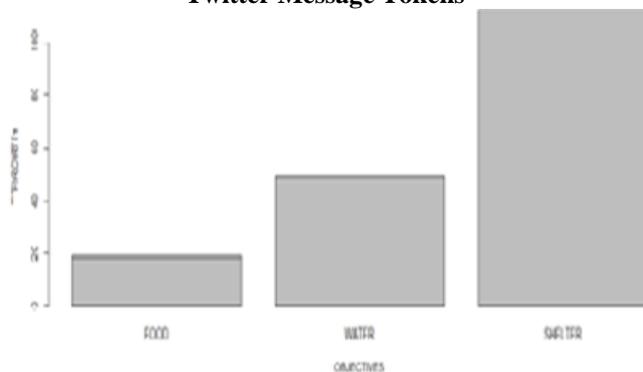


Fig 5. Each Prioritization Objects the Frequency of flood Twitter Messages

V. PERFORMANCE ANALYSIS

Prioritization of Key Objectives during Floods is much better than current existing methods used to help natural disaster affected people. The existing methods are not give more relevant prioritizations which aren't enough to needy people, these include methods like boots on the ground survey or following by the book protocol which isn't dynamic in any aspect. To overcome these disadvantages, in our work aims to improve upon these methods or more appropriately provide a new method to handle such cases all together. If the flood relief group can acquire this type of information they can quickly focus and address the issue which is heal as most important by the people, which is the more optimizing way of helping needy people those are affecting by natural disaster. This work is also flexible as it

offers the users to prefer the size of sample space would like to analyse for improving the accuracy of evaluate as well as effort provide lodgings to needy people with in the short span of time.

VI. CONCLUSION

With this analysis and growing technology around disaster aid relief, such as project loon, and many others, we can conclude the requirements or objectives needed by the victims of floods much faster, with greater efficiency and more precision than any other existing method it completely eliminates any unnecessary hassle and streamlines the process of feedback and suggests a future course which adds precious time to be able to do more and to react to growing situations before they spiral out of control and ensure the minimum public to disaster relief, organization latency.

FUTURE ENHANCEMENT

As of now the technology and infrastructure built around calamity relief is only efficient when it is concerned with floods and similar natural disasters, but as these technologies become more effective in handling other calamities, our work can be seamlessly integrated to focus on any natural disaster.

REFERENCES

- Brooks, Nick, and W. Neil Adger. "Country level risk measures of climate-related natural disasters and implications for adaptation to climate change Nick Brooks and W. Neil Adger Tyndall Centre for Climate Change." (2013).
- Datar, Ashlesha, Jenny Liu, Sebastian Linnemayr, and Chad Stecher. "The impact of natural disasters on child health and investments in rural India." *Social Science & Medicine* 76 (2013): 83-91.
- Chung, D.S., Nah, S.: Media credibility and journalistic role conceptions: views on citizen and professional journalists among citizen contributors. *J. Mass Media Ethics* 28(4), 271–288 (2013)
- Li, R., Lei, K.H., Khadiwala, R., Chang, K.-C.: TEDAS: a Twitter-based event detection and analysis system. In: *Proceedings of 28th International Conference on Data Engineering*, pp. 1273–1276 (2012)
- Lingad, J., Karimi, S., Yin, J.: Location extraction from disaster-related microblogs. In: *Proceedings of the 22nd International World Wide Web Conference Companion*, pp. 1017–1020 (2013).
- Unankard, S., Li, X., Sharaf, M., Zhong, J., Li, X.: Predicting elections from social networks based on sub-event detection and sentiment analysis. In: Benatallah, B., Bestavros, A., Manolopoulos, Y., Vakali, A., Zhang, Y. (eds.) *WISE 2014. LNCS*, vol. 8787, pp. 1–16. Springer, Heidelberg (2014). doi:10.1007/978-3-319-11746-1 1.
- M. Hu and B. Liu, "Mining and Summarizing Customer Reviews," in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2004, pp. 168–177.
- Sakaki, T., Okazaki, M., Matsuo, Y.: Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Trans. Knowl. Data Eng.* 25(4), 919–931 (2013).
- X. Ding, B. Liu, and P. S. Yu, "A Holistic Lexicon-based Approach to Opinion Mining," in *Proceedings of the 2008 International Conference on Web Search and Data Mining*, 2008, pp. 231–240.
- M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-Based Methods for Sentiment Analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.
- M. Z. Asghar, A. Khan, S. Ahmad, M. Qasim, and I. A. Khan, "Lexicon-enhanced sentiment analysis framework using rule-based classification scheme," *PLOS ONE*, vol. 12, no. 2, pp. 1–22, 2017.
- Z. Jin, Y. Yang, X. Bao and B. Huang, "Combining user-based and global lexicon features for sentiment analysis in twitter," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 4525-4532.

Prioritization of Key Objectives During Floods

13. Turney, P.D., Littman, and M.L.: Measuring Praise and Criticism: Inference of Semantic Orientation from Association. *ACM Transactions on Information Systems* (2003) 315–346
14. Kamps J., Marx, M., Mokken, R.J.: Using WordNet to Measure Semantic Orientation of Adjectives. *IREC vol. IV* (2004) 1115–1118.
15. <https://x.company/loon/>
16. <http://www.southoxon.gov.uk/services-and-advice/environment/severe-weather/flooding/flood-plans-0>

AUTHOR PROFILE



Prof. Vijayalakshmi Kakulapati received **Ph.D in Computer Science & Engineering** in the area of Information Retrieval from JNTU Hyderabad. She is having 25 years of industry and teaching experience.. She is a member of various professional bodies like **IEEE, ACM, CSTA, LMISTE, LMCSI, IACSIT, FIETE** and few more.

She is serving as an Editorial Board member to various International and National journals including *Journal of American Association of Science and Technology (AASCIT)*.

Apart from these she is continuously serving many National & International Conferences as a review member and various International Journals including *IJCT, Journal of Big Data, oriental journal of computer science and Egyptian Informatics Journal*. She is having more than 70+ publications in National, International Journals and conferences. Out of these 20 are Springer Chapters, 4 ACM, 8 IEEE and 2 in Elsevier. She took active role in National Conferences and four International Conferences as Session chair including ICACCI 2018. She has been received an **Excellence in Research award from INDUS foundation, Innovative research and dedicated academician award from global awards, 2018, Best Senior Researcher award from GECL awards and appreciation awards for high impact factor publication and conducting e-learning programs**. She gave 16+ guest Lectures on “Information Retrieval Systems, Advance data mining and Big data analytics and health informatics” in various engineering colleges. she organized various Workshops include AICTE sponsored and DST sponsored workshops. Currently she is working with **big data analytics**, health informatics, Data science and Internet of Things. She is also working as a coordinator on various programs like IUCEE, NPTEL Globarena etc. She invited as a key note speaker for her research in Health Informatics from Brazil, china and other countries.