

## Analyzing the Impact of Online News Papers using Readers Behaviors

\* J.Ragini<sup>1</sup>, Dr.P.Tamilselvi<sup>2</sup>.

<sup>1,2</sup>Department of Computer Application, Jeppiaar Engineering College, Chennai-119, Tamilnadu, India

\*Corresponding Author: Email: ragini.j18@gmail.com

### ABSTRACT

The internet is by and large employed for information retrieval on various topics now a day. The information deduction on different areas is still growing. The readers' comments and feedbacks are playing vital role in reader's interest. In this paper, various online newspapers are considered for evaluating the highly reachable paper among public. The news paper's popularity is predicted based on the Reader's Behaviors. For making the valuation, the reader's tweets counts and favorites are mainly considered along with other Reader's Behaviors factors such as followers and following factors. To analyze the highly prepared newspapers by the readers are measured by calculating the ratio and percentage among the Reader's behaviors factors. The analysis is mainly made for measuring the impact of the papers towards the society. From the result of the analysis it is apparent that online newspapers play a crucial role in the spread of information on that site, and that network structure affect dynamics of information flow.

**Keywords:** online newspapers, twitter, reader's behavior analysis, reader's comments, ranking, prediction.

### INTRODUCTION

Social websites are platforms that allow common persons to create and issue contents. The worldwide popular the people website, Twitter, reader's act as a huge detector force by reporting events in real time. There is other specialized social websites that are attentive on activity, sports instruction, business and political interaction. Since there are many reader's sharing their opinions and experiences via social media. Social websites may promote news stories which are expected to make greater impact to attract more readers.

Reader's behavior analysis can be used to takeout mass opinions from such discussions. A basic task in reader's behavior analysis is classifying the polarity of a given text at the document, sentence, or categorization level — whether the expressed opinion in an article, an opinion or a behavior categorization is positive, negative, or neutral. Hard, "accurate polarity" reader's behavior categorization looks, for instance, at emotional state such as "angry," "sad," and "happy." Sometimes, the structure of reader's behavior and topics is fairly complex. Also, the problem of reader's behavior analysis is non-monotonic in respect to sentence extension and stop-word substitution. To address this issue a number of rule-based and reasoning-based approaches have been applied to reader's behavior analysis, including defensible logic programming also, at this time a number of tree traversal rules concern to syntactic parse tree to extort the topicality of behaviors in open domain setting.

Open resource software tools categorize machine learning, financial, and typical language processing technique to computerize reader's behavior investigation on large collection of texts, together with web pages, online information, internet conversation groups, online reviews, web blogs, and open media. Knowledge-based systems, instead, expand use of frequently accessible resources, For example, WordNet-Affect, Senti Word Net, and Sentic Net, to extort the semantic and affective information related with expected language concepts. Reader's behavior assessment can also be perform on image domestic i.e. images and videos. Individuality effectively approach in this direction is Senti Bank utilizing an adjective noun pair demonstration of visual content. This allows movement to a more problematical understanding of reader's behavior based on an 11 point scale. Or else, text can be isolate reader's behavior strength score if the goal is to determine the reader's behavior in a text logically than the essentially popularity and disturbance of the text. The paper is classify as follow: the related works about the Reader's Behaviors analysis is discuss in section II, the proposed system for news articles analysis is describe in section III, the outcomes are discuss in section IV, and the conclusion in section V.

### RELATED WORKS

Kristina Lerman, 2010, proposed the data sets offers a rich source of evidence for studying dynamics of individual and group behavior, the strike of networks and global patterns of the flow of information. Social networks play a important role in the spread of information on these sites, and that networks structure affects dynamics of information

flow. Jong Gun Lee, 2010, was proposed a different approach rooted in survival analysis. Where forecast the accurate lifetime of an individual is very hard and almost impossible but predicting the likelihood of one's survival longer than a threshold, or another individual is possible. Manos Tsagkias, 2010, was proposed a report on solid performance for predicting new comment volume in the long run, after short evaluation. The number of user supplied comments on a news article may be indicative of its importance, interesting, & impacting.

Chiao-Fung Hsu, 2009, was proposed by Chiao-Fang Hsu et al., where estimate a machine learning-based approach for ranking comments on the Social web based on the community's expressed preferences. Which can be used to support high-quality comments and filter out low-quality comments. Gabor Szabo, 2008, was proposed that predictions are more accurate for submissions for which attention decays quickly. Whereas predictions for evergreen content will be prone to large errors, which could further impact the popularity and smoother the decay of interest over time.

## SYSTEM DESIGN

In this system, we evaluate the Reader's Behaviors analysis that can be measured by the popularity contents and the factors are chosen to impact the popularity (e.g. Total number of tweet, the total number of favorites). The system process is given as a flow diagram in Fig 1.

**A. Ranking based on ratio:** We use as factors, the comments that are visible to the reader's, such as the total number of tweet and the total number of favorites. The ratio can be calculated by these two factors. A ratio is a comparison between two numbers and it is described as a fraction. If P and Q are any two numbers, then the ratio of P and Q is P/Q.

The ratio among the tweet and the favorite is calculated by,

$$R_i = \frac{P_i}{\text{gcd}(P_i, Q_i)} : \frac{Q_i}{\text{gcd}(P_i, Q_i)}$$

Where,

$R_i$  - ratio of tweet and favorites  $P_i$  – number of tweets  $Q_i$  -number of favorites

GCD – Greatest common difference

**B. Ranking based on percentage:** The tweet values are very high, for the convenient process, the values of all the new papers are divided by 100 and the resulted value is considered for the analysis. To ensure the results what derived from the ratio, the percentage is calculated between the newly derived tweet values and the favorite values. The percentage is calculated by,

$$P_i = \frac{T_i/100}{F_i} \times 100$$

$P_i$  – Percentage of tweet and favorites

$T_i$ – number of tweets

$F_i$  – number of favorites

The Readers Behaviors analysis is done based on the basic factors of reader's behaviors such as tweets, favorites, followers and following etc. The processing steps are given as an algorithm below.

### Algorithm:

1. Collect the Readers Behaviors factors.
2. Collect the reader's behaviors factors of various newspapers.
3. Calculate the ratio and percentage using the two major factors such as tweet and favorites.
4. Evaluate the ranking based on the ratio.
5. Compare the ratio values for finding the highly reaching newspapers.
6. Stop

## RESULTS AND DISCUSSION

The newspapers are analyzed based on the two major factors such as tweet and favorites. Using internet, various online newspapers information are collected to measure the high reachability among the public. The sample values from the collected data are shown in the table Table.1.

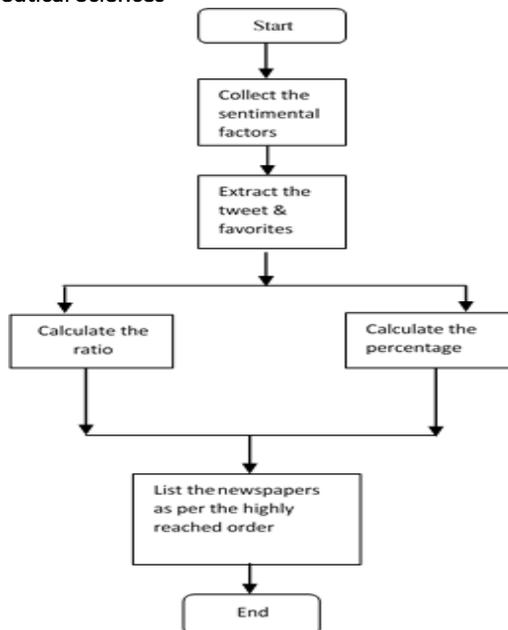


Fig.1.Flow diagram for ranking based on user's performance

Table.1.Summary of the data sets analyzed in this paper

News papers	Tweet	Followers	Following	Favorites(likes)
Times of India	172K	3.97M	464	176
Times Now	76.2K	2.06M	103	79
The Hindu	101K	1.36M	151	48
The Economics Times	139K	606K	30	118
Indian Express	121K	762K	646	348

The other factors such as the followers and followings are also collected. But, the main factors tweet and favorites are dominating the previous factors. Because of this, the previous factors not considered for the analysis. The reachability of a newspaper is evaluated based on the values of the tweet in the tabular column is shown as bar chart below.

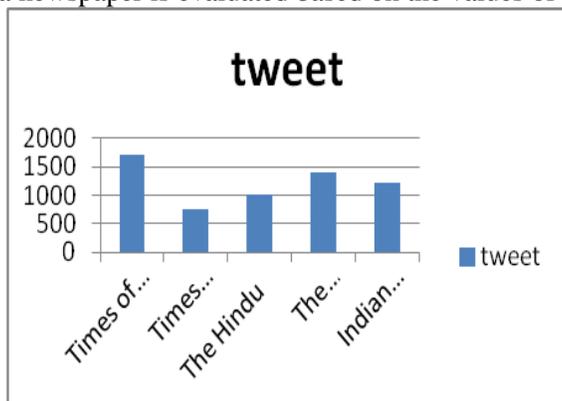


Fig.2.Comparison based on Tweet factor

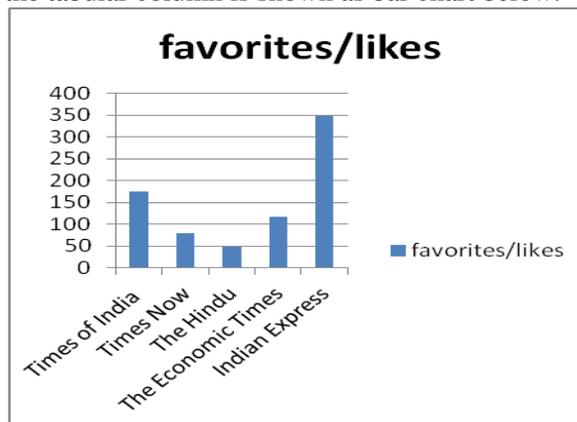


Fig 3. Comparison based on 'like' factor

The values of the favorites in tabular column is also represented as bar chart. Both these bar chart, symbolize the status of newspapers before Ranking. Once Ranking is performed, the status of news articles is described in a final bar chart diagram. This is represented below.

The tweet values are compared in Fig 2, the favorite values are compared in Fig 3 and the ratio which is used for measuring the performance evaluation is shown in Fig 4.

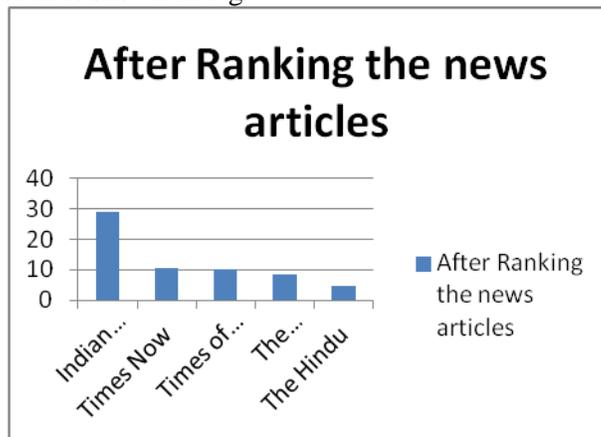


Fig 4. Comparison based on rank values

## CONCLUSION

The paper is projecting the highly reaching newspapers to the public. To evaluate the newspapers reachability, the factors which are selected by the readers are not taken. The factors which are measured automatically by the system such as tweet count, favorites, followers and following are considered. There is less possibility of wrong data which are fed by the human. From the result, it is very obvious that Indian Express is prepared by many readers in online. The Times Now and Times of India are the next highest reaching newspapers among the people. From the analysis, The Hindu has got the fifth place to reach the readers.

## REFERENCES

- Alexandru Tatar, Panayotis Antoniadis, Marcelo Dias De Amorium, Serge Fdida. From Popularity Prediction to Ranking Online News. *Social Networks Analysis and Mining*, Springer, 4, 2014, 174.
- Hsu C, Khabiri E, Caverlee J, Ranking comments on the social web. In: *Computational Science and Engineering*, 2009. CSE'09. International Conference on, IEEE, 4, 2009, 90-97.
- Lee J, Moon S, Salamatian K, An approach to model and predict the popularity of online contents with explanatory factors. In: *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 01, 2010.
- Lerman k, Ghosh R, Information Contagion: an Empricial Study of the Spread of News on Digg and Twitter Social Networks. In: *Proceedings of 4th International Conference on Weblogs and Social Media (ICWSM)*, 2014.
- Szabo G, Huberman BA, Predicting the popularity of online content, *Commutations of the ACM*, 53(8), 2008, 80
- Tsagkias M, Weerkamp W, De Rijke M, News comments: Exploring, modeling, and online prediction. In: *Proceedings of the 32nd European conference on Advances in Information Retrieval*, Springer, ECIR2010