International Journal of Scientific Engineering and Research (IJSER)
www.ijser.in
ISSN (Online): 2347-3878
Volume 1 Issue 2, October 2013

Ranking Assessment of Event Tweets for Credibility

Sravan Kumar G

Student, Computer Science in CVR College of Engineering, JNTUH, Hyderabad, India

Abstract: ***Online social network services have become a popular web activity to establish online social relationship among the people all around the world. These online social networking services allow its users to share opinions or posts on any high impact events in the world. The primary task of these services is to sort out credible posts and provide credible information about the event to the users. In this paper, the focus has been on Twitter, a rapidly growing micro blogging platform, which provides a large amount, diversity and varying quality of content. As Twitter is open to all, it emerged as an excellent means to disseminate information to a large user community in the shortest time. Due to very open, uncontrolled nature, Twitter has become vulnerable to incredible information from malicious and credulous users. Consequently, it is important to formulate sophisticated methods for analysis of credibility and relevance for ranking tweets. In this paper, tweets of an event posted by users have been collected and allowed to perform annotation process on those tweets by three human annotators to assess the tweet credibility. In order to provide ranks to the tweets according to the features, content based features are extracted. The performance of the ranking strategy used to rank tweets according content based features has been enhanced by the re-ranking strategy which uses context specific features such as event specific words to re-rank the tweets. After re-ranking tweets, an evaluation has been carried out by NDCG metric to measure the accuracy of re-ranking strategy.***

Keywords: Tweets, Human Annotation, TF, IDF, PRF, BM25, NDCG

1. Introduction

Online social networking services like Twitter, Facebook and MySpace have emerged as popular media for information sharing. Users using these services are constantly increasing. Web users keep up with the latest information through popular online social services. With the evolution of these online social networking services, two changes occurred in the usage of internet [1]. Firstly, the internet replaced traditional media like television and print media as a source for obtaining news and information about current events. Secondly, the internet has provided platform for common people to share information and express their opinions.

Quick response time and high connectivity speed of internet made the users on online social services to disseminate the in-formation or news quickly with in fraction of seconds. Dissemination of news or information through traditional media is credible where source of information are few and known. Due to the anonymous and unmonitored nature of internet, a lot of content generated by many sources of information may be credible or incredible [1]. Among all online social networking sites, Twitter is fastest growing social networking site that provides a micro-blogging service to users where user can post their messages called tweets. Twitter emerged as major news source and information dissemination agent over last few years. Twit-ter is a crowd-sourced medium where users on it may generate credible or incredible information about event [1]. Fake news and rumors also propagate along with genuine news. It is difficult to identify the credible tweets during high impact events manually. In this research work, a ranking scheme is proposed to present the user a ranked output of tweets according to the credibility of information in the tweet.

2. Review of Literature

Online social networking sites such as Facebook, LinkedIn and Twitter allow users to meet new people, establish professional connections and more [9]. Twitter provides a micro blogging service where users can send short messages called tweets that appear on their friend's page. A user on Twitter is uniquely identified by his username and optionally by his real name. A Twitter user can start following another user X. Consequently, that user receives user X's tweets on his/her own page [9]. Tweets can be grouped by hash tags which are popular words, beginning with a "#" character [9]. A user can decide to protect her profile. By doing so, any user who wants to follow that private user needs her permission.

Twitter has recently merged as a popular social system where users share and discuss about everything, including news about events [2]. With simple interface only 140 character messages can be posted. Twitter is increasingly becoming a system for obtaining real time information and source for news and latest trends [2]. Twitter emerged as an excellent means to disseminate information to a large user community in the shortest time [11]. On the negative side, this very open uncontrolled nature of twitter service makes micro blogging vulnerable to false information from malicious users [11]. Consequently, it is important formulate sophisticated methods for analysis of relevance and trustworthiness for ranking tweets. Ranking that considers content based features, place the most credible and popular tweets in the top slots.

2.1 Acronyms

Table 1: Acronyms

Term	*Definition*
PRF	Pseudo Relevance Feedback.
NDCG	Normalized Discounted Cumulative Gain.
BM25 Metric	Best-Match 25 Metric is a Text Similarity Metric.
IDF	Inverse Document Frequency.

2.2 Definitions and Background

- **Tweet:** Twitter provides a micro blogging service to users where users can post their message or statuses called Tweets. Each tweet is limited to 140 characters and allows http links to be included in it [9].

- **Credibility:** Credibility is defined as the quality of being trusted and believed in. A tweet is said to contain credible information about news event, if we trust or believe that information in the tweet is said to be correct/true [1].

- **Human Annotation:** It is procedure which takes the help of three annotators to establish ground truth regarding the presence of credible information in tweets related to news event [1].

- **BM25:** It is a ranking function used by search engines to rank matching documents according to their relevance to a given search query. BM25 is a retrieval function that ranks a set of documents based on the query terms appearing in each document, regardless of the inter-relationship between the query terms within a document (e.g., their relative proximity) [7]

- **Inverse Document Frequency:** It is a statistical popular measure of a word's importance. It is defined as the logarithmic ratio of the number of documents in a collection to the number of documents containing the given word. This means rare words have high IDF and common words have low IDF [10].

- **NDCG:** Normalized Discounted Cumulative Gain (NDCG) which is a family of ranking measures widely used in applications. It is popular measure for evaluating web search and related tasks [10]. NDCG has two advantages compared to many other measures. First, NDCG allows each retrieved document has graded relevance while most traditional ranking measures only allow binary relevance. That is, each document is viewed as either relevant or not relevant by previous ranking measures; while there can be degrees of relevancy for documents in NDCG. Second, NDCG involves a discount function over the rank while many other measures uniformly weight all positions [10]. This feature is particularly important for search engines as users care top ranked documents much more than others [8].

3. Existing System

A task of social network users is to decide whose updates to subscribe in order to maximize the relevance, credibility and quality of information received. Previous research explored the credibility on Twitter with respect to trending topics. Credibility of a topic on Twitter may not be a good indicator to judge the credibility of the content of the tweet. Thus, assessment techniques must be required at the atomic level of the information on twitter, i.e. at a tweet level.

4. Proposed System

In this paper, credibility is assessed at tweet level by considering content and context specific based features that are used to rank the tweets according to the credible information contained in tweets. Consequently, users on Twitter who are in need to know the credible information on particular event or topic will be shown according to the level of credibility in the tweets.

5. Methodology and Analysis

With the evolution of online social networking and micro-blogging mediums, two major changes have occurred in the landscape of the Internet usage - firstly, the Internet is re-placing traditional media like television and print media as a source for obtaining news and information about current events; secondly, the Internet has provided a platform for common people to share information and express their opinions [1]. One major difference between dissemination of news or information through traditional media and Twitter is that, Twitter is a crowd-sourced medium [1]. In contrast to television, print and news websites, the source of information are few and known.

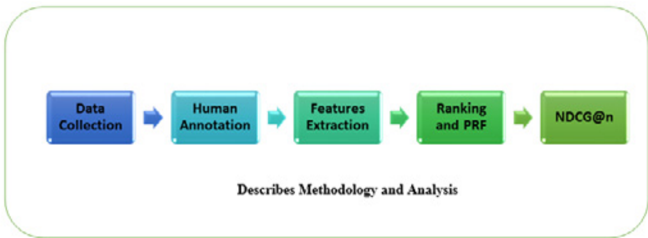


Figure 1: Methodology and Analysis

Due to the anonymous and unmonitored nature of the Internet, a lot of content generated on Twitter maybe incredible and it is hard to identify the tweets with credible information manually [1]. Proposed solution provides automated ranking scheme to present the user a ranked output of tweets according to the credibility of information in the tweet.

The proposed solution includes the following modules:

1) **Data Collection:** The proposed solution provides interface to users to tweet on particular event which is posted by admin. All tweets on event are collected by admin are provided to human annotators.

2) **Human Annotation:** Tweets by all users are assessed semantically by human annotators. For each tweet, human annotators are provided by three options from which annotator have to select one of the options. Three options are:

   - Credible
   - Incredible
   - Not Relevant

   We considered the tweet and forwarded to next module only if that tweet is provided by credible option by at least one human annotator.

3) **Feature Extraction:** All the annotated tweets by human annotator are provided to admin to extract the features of annotated tweets. The extracted features are length of the tweet and number of unique words.

4) **Ranking and Re-ranking:** Features that are extracted from the annotated tweets are considered to rank the annotated tweets. In ranking, we used the feature, number of unique words as a measure of the information richness of a tweet. Intuitively, a tweet containing more number of unique words is apt to contain more information than a short. Hence, we ranked the tweets according to the feature called number of unique words in the tweet.

PRF ranking which is also known as re-ranking uses Inverse Document Frequency of event related words to calculate BM25 metric value which is text similarity metric. Based on BM25 value, tweets are re-ranked.

BM25: It is a ranking function used by search engines to rank matching documents according to their relevance to a given search query. BM25 is a retrieval function that ranks a set of documents based on the query terms appearing in each document, regardless of the inter-relationship between the query terms within a document (e.g., their relative proximity). Given a query Q, containing keywords [q1, q2… qn] and then BM25 score of the Tweet D is [7].

$$\text{score}(D,Q) = \sum_{i=1}^{n} \text{IDF}(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot (1 - b + b \cdot \frac{|D|}{\text{avgdl}})}$$

Here f (qi,D) is qi’s term frequency in the tweet D. |D| is the length of the tweet in words. avgdl is the average length of the tweet. k1 and b are free parameters. Concerning the internal parameters, the model provides no guidance on how these should be set. This may be regarded as a limitation of the model. However, it provides an opportunity for optimization, given some evaluated set of queries and relevance judgments in the traditional retrieval experiment style. A significant number of such experiments have been done, and suggest that in general values such as $0.5 < b < 0.8$ and $1.2 < k1 < 2$ are reasonably good in many circumstances. In this research, parameter values are k1=1.2 and b=0.75.

Inverse Document Frequency: is a statistical popular measure of a word’s importance. It is defined as the logarithmic ratio of the number of documents in a collection to the number of documents containing the given word. This means rare words have high IDF and common words have low IDF [10].

$$\text{IDF}(q_i) = \log \frac{N - n(q_i) + 0.5}{n(q_i) + 0.5}$$

n (qi) is number of tweets containing qi, N is number of tweets.

Evaluating Ranking using Metric: Tweets are ranked based on BM25 value are evaluated using NDCG metric. They provide four options for each tweet. Selecting one of the options for all tweets, NDCG gives the percentage of efficiency of the ranking that used to rank tweets.

NDCG: Normalized Discounted Cumulative Gain (NDCG) which is a family of ranking measures widely used in applications. It is popular measure for evaluating web search and related tasks [10].

$$DCG_p = \sum_{i=1}^{p} \frac{2^{rel_i} - 1}{log(1+i)}$$

Let reli be the judgment of i^{th} tweet (4=Credible, 3=May be Credible, 2=Incredible, 1=Not Relevant). The normalized DCGp is the DCGp divided by the DCG of the ideal (Maximum DCG value).

Normalized Cumulative Gain (NDCG) at rank n:

Normalize DCG at rank n by the DCG value at rank n of the ideal ranking. The ideal ranking would first return the documents with the highest relevance level, then the next highest relevance level, etc,

6. Experimental Results

6.1 Data Collection

Figure 2: Interface to Post Tweet for Admin

Figure 3: Interface to View Event for Admin

The above Figure 2 and 3 show pages which are responsible for both posting and viewing events. Activities like posting and deleting events will be allowed by administrator only. In this paper, an event by name "Issue of Hyderabad after announcing Telangana State is posted by administrator". The Event posted by administrator will be seen by the users who are registered. Posted Event gives clear description about the event and also image related to that context. The description given for this event is – "Opinions of Andhra Pradesh people on the issue of Hyderabad during initiation of Telangana state". The image used in this context is a map of Andhra Pradesh state that shows bifurcation of Telangana state.

Figure 4: Interface to View Event for User

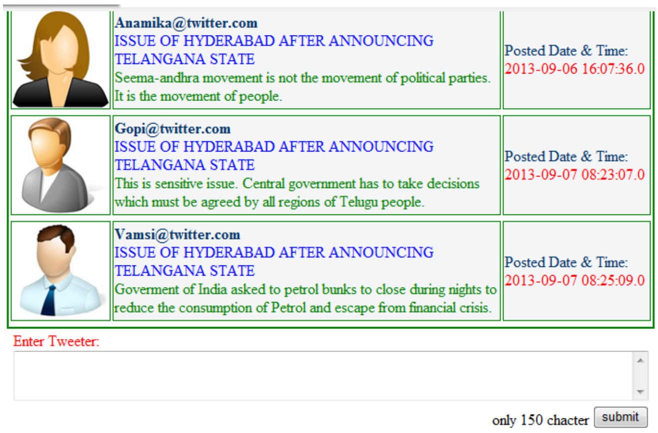


Figure 5: Interface to Post Tweet for User

The above figures 4 & 5 display the interface to view and post tweets on the event posted by admin. User can view tweets of different users of the same event. The above figure 5 shows the tweet of a user followed by posted date along with time. User can post any number of tweets on single event posted by administrator.

6.2 Human Annotation

The following Figure 6 shows interface by which the human annotators annotate the tweets by selecting one of the three options. Human Annotators who are well known of event will eliminate the tweets which are not related to the event posted by admin. In this paper, three annotators perform annotations process to give annotated tweets to the admin. After the process of Annotation, Annotators provide the tweets which are related to the event and thus annotated tweets collected by admin to perform further processing. In this paper, human annotators annotated 102 tweets of an event and provided 57 annotated tweets of an event for ranking according to the credibility of information contained in it. In this project, Human Annotation process eliminated 44% of tweets which are not related to event.

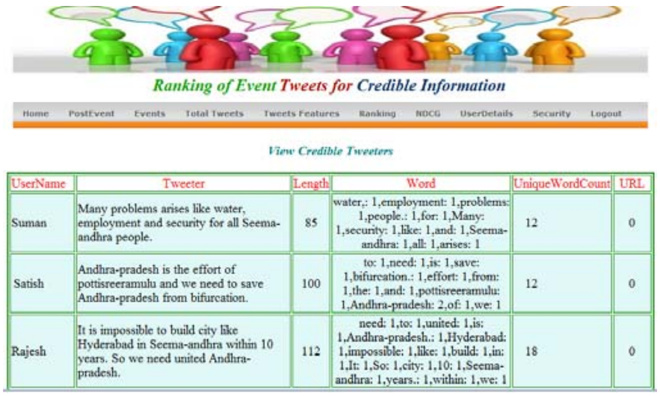
Figure 6: Interface to Annotate Tweets for Annotators

6.3 Feature Extraction

The Figure 7 shown below shows message based features such as length of the tweet, unique words count and URL

International Journal of Scientific Engineering and Research (IJSER)
www.ijser.in
ISSN (Online): 2347-3878
Volume 1 Issue 2, October 2013

count. Length of the tweet represents number of characters present in the particular tweet and unique word count represents number of unique words present in the corresponding tweet.



UserName	Tweeter	Length	Word	UniqueWordCount	URL
Suman	Many problems arises like water, employment and security for all Seema-andhra people.	85	water,: 1,employment: 1,problems: 1,people.: 1,for: 1,Many: 1,security: 1,like: 1,and: 1,Seema-andhra: 1,all: 1,arises: 1	12	0
Satish	Andhra-pradesh is the effort of pottisreeramulu and we need to save Andhra-pradesh from bifurcation.	100	to: 1,need: 1,is: 1,save: 1,bifurcation.: 1,effort: 1,from: 1,the: 1,and: 1,pottisreeramulu: 1,Andhra-pradesh: 2,of: 1,we: 1	12	0
Rajesh	It is impossible to build city like Hyderabad in Seema-andhra within 10 years. So we need united Andhra-pradesh.	112	need: 1,to: 1,united: 1,is: 1,Andhra-pradesh.: 1,Hyderabad: 1,impossible: 1,like: 1,build: 1,in: 1,It: 1,So: 1,city: 1,10: 1,Seema-andhra: 1,years.: 1,within: 1,we: 1	18	0

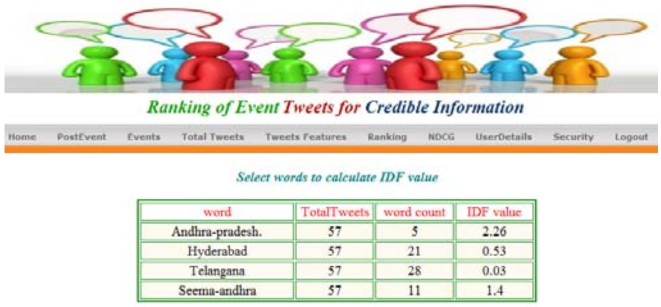
Figure 7: Feature Extraction

6.4 Ranking and Re-Ranking

Tweets shown in Figure 8 corresponded with rank which is calculated based on the features called message based features. The Tweet which contains more number of unique words, placed at rank1 slot. Number of unique words feature is the measure of the information richness of the tweet and long sentence with more number of unique words contain more information. Re-Ranking technique which is called as Pseudo Relevance Feedback (PRF) is used to enhance the performance of ranking results. In PRF technique, we used BM25 metric which calculates the text similarity between tweet and the query set which contains most frequent or event related words. BM25 metric employs Inverse Document Frequency (IDF) value of each query set words.

Rank	Tweeter
Rank1	Police force are providing security at both region borders. Employees must show the ID card with photograph to get enter in stadium.
Rank2	To make Rahul as PM of India in 2014 elections, Sonia divided AP. She's in a process merging TRS in to congress to get power there.
Rank3	This is sensitive issue. Central government has to take decisions which must be agreed by all regions of Telugu people.
Rank4	Some political parties are requesting central government to make Hyderabad as union Territory for which Telangana people never accept.
Rank5	It's not that Telangana is only under developed. There are many places in North Andhra and Rayalaseema which are at worst.

Figure 8: Ranked Tweets



word	TotalTweets	word count	IDF value
Andhra-pradesh.	57	5	2.26
Hyderabad	57	21	0.53
Telangana	57	28	0.03
Seema-andhra	57	11	1.4

Figure 9: Displaying IDF values of Event Specific Words

The above Figure 9 shows the IDF values for the words which are related to the event. IDF is a popular measure of a word's importance. It is defined as the logarithm of the ratio of number of tweets in corpus to the number of tweets containing the given word. The word which appear rarely in all the tweets possess high IDF value and the word which occurs frequently in tweets corpus contain low IDF value.

Tweets	BM25 Value	Rank
Hyderabad was developed during Nizam time. We have drainage line system before formation of Andhra-pradesh.	4.72	Rank1
It is impossible to build city like Hyderabad in Seema-andhra within 10 years. So we need united Andhra-pradesh.	4.19	Rank2
It is unfair to divide Telugu people according to their regions. People are demanding united Andhra-pradesh.	3.82	Rank3
We trusted Hyderabad belongs to all and we invested in it by leaving Seema-andhra cities without developing.	3.27	Rank4
People of Telangana never received fruits from Andhra-pradesh. So we demand separate Telangana with Hyderabad as capital.	2.82	Rank5
We Seema-andhra people were shifted our capital and high court from Kurnool and Guntur to Hyderabad.	2.37	Rank6

Figure 10: Re-Ranked Tweets

6.5 Evaluating Ranking using Metric

S.no	Tweets	Select option
1	Hyderabad was developed during Nizam time. We have drainage line system before formation of Andhra-pradesh.	Credible May be Credible In Credible Not Relevant
2	It is impossible to build city like Hyderabad in Seema-andhra within 10 years. So we need united Andhra-pradesh.	Credible May be Credible In Credible Not Relevant
3	It is unfair to divide Telugu people according to their regions. People are demanding united Andhra-pradesh.	Credible May be Credible In Credible Not Relevant

Figure 11: Interface to Evaluate Rankings

In the above Figure 11, we evaluated ranking and re-ranking by using NDCG metric on top three tweets.

ISSN (Online): 2347-3878
Volume 1 Issue 2, October 2013

Figure 12: Displaying NDCG@3 value

In the above Figure 12, maximum DCG@3 value and NDCG@3 value is shown. NDCG@3 value is 0.84208 which represents that ranking used was 84.2% efficient.

7. Conclusion and Future Work

For a successful application of the proposed solution, tweets of high impact events must be ranked according to the credible information contained in them. In this paper, tweets of users, posted on event are collected to perform annotation process by three human annotators to obtain credible tweets. After obtaining credible tweets of an event, message features of those tweets are extracted and ranked according to message features and relevancy.

Moreover, ranking evaluation metric has been used to evaluate re-ranking process. Thus, the proposed solution supports the administrator to provide the tweets of an event that contains credible information to the users who require credible information of an event. The work, presented in this paper introduces possible direction for further work. The proposed solution can be enhanced by eliminating the Human Annotation process that requires experts to annotate the tweets of events. Human Annotation process would be tedious task to annotators and consumes more time when the large volume of content i.e., Tweets are posted on event.

The limitation of Human Annotation process which establishes the ground truth can be overcome by developing self-learning mechanism and automatically adopting systems that do not require human annotators to annotate tweets manually. Finally, integration of the suggested research approach along with the proposed approach would be more effective to rank the large volume of tweets of high impact events.

8. Acknowledgment

I would like to thank and express deep sense of gratitude to Mr. A Seetharam Nagesh, Assistant Professor at CVR College of Engineering and Dr. N V Rao, Professor in CSE Department, CVR College of Engineering, for their valuable time and useful suggestions.

References

[1] Aditi Gupta, Ponnurangam Kumaraguru. “Credibility Ranking of Tweets during High Impact Events”, ACM, 2012
[2] F Benevenuto, G Magno, T Rodrigues, and V Almeida, “Detecting Spammers on Twitter”, In CEAS, 2010
[3] K R. Canini, B Suh, and P L Pirolli, “Finding Credible Information Sources in Social Networks based on Content and Social structure”, SocialCom, 2011
[4] C Castillo, M Mendoza, and B Poblete, “Information Credibility on Twitter”, WWW 2011 pages {675-684}, ACM, 2011
[5] Chen, R Nairn, L Nelson, M Bernstein, and E Chi. “Short and Tweet: Experiments on Recommending Content from Informa-tion Streams”. CHI '10, pages {1185-1194}, ACM, 2010.
[6] S Chhabra, A Aggarwal, F Benevenuto, and P Kumaraguru. “Phi.sh/$ocial: The Phishing Landscape through Short urls”. CEAS 2011, pages {92-101}, ACM, 2011.
[7] Y Duan, L Jiang, T Qin, M Zhou, and H. Y Shum, “An Empirical Study on Learning to Rank of Tweets”. Proceedings on 23rd International Conference on COLING, 2010, pages {295-303}
[8] Yining Wang, Liwei Wang, Di He, Wei Chen. “A theoretical Analysis of NDCG Ranking Measures”. Workshop and Conference Proceedings Vol (2013)
[9] M. McCord, M Chuah. “Spam Detection on Twitter using Traditional Classifiers”. IEEE 2011
[10] Kishore Papineni. “Why Inverse Document”, IBM T.J. Watson Research Centre, NY, USA
[11] Srijith Ravi kumar, Raju Balakrishnan and Subbarao,”Ranking Tweets considering Trust and Relevance”, ACM 2010

Author Profile

Sravan Kumar G received his B.Tech Degree in Information Technology from Sphoorthy Engineering College, JNTUH, Hyderabad, India in 2010. Now he is pursuing M.Tech in Computer Science from CVR College of Engineering, JNTUH, Hyderabad, India.