

# Twitter News Credibility Meter

Zareen Sharf  
PhD Scholar, SZABIST  
Islamabad, Pakistan

Anwar us Saeed, Ph.D  
Professor Data Mining, SZABIST  
Islamabad Pakistan

## ABSTRACT

Online users are usually naïve and tend to believe everything they come across on social media. Micro blogging is gaining more and more popularity as a journalism tool with the passage of time and can create extraordinary hype about events that may or may not be true. It therefore becomes a social responsibility to devise mechanisms that might authenticate and validate the credibility of a news event or crisis situation being posted on social media sites. Twitter being one of the most popular real time news sources also becomes one of the most powerful news emanating mediums. It has known to cause considerable damage by spreading rumors in the past. This study is therefore dedicated for devising an application that might help in establishing a certain level of accuracy regarding the information posted on twitter. The application performs a keyword based retrieval of tweets and then assesses validity based on predefined criteria.

## General Terms

Data mining, journalism through social media, Social Networks news broadcast.

## Keywords

Credibility, News events, Twitter, Social Media

## 1. INTRODUCTION

Internet usage has been radically altered due to extensive utilization of social networking & micro blogging. Not only these services have become prominent sources for dissemination of news and current event but also provide a platform for information sharing and freedom of expression to the general masses. One of the major reasons for this shift of news medium is the immediate propagation of potential news and current events. The focus of our study will be mainly on twitter. Twitter has become one of the most widely used micro blogging services for the diffusion of mega events ranging from a crisis situation to an entertainment scoop Not only the tweets contain volumes of data that can provide valuable information but parameters like geographical location, major interests, number of followers and frequency of tweets can also make significant contribution in the various studies being conducted at the time.

Since twitter is accessible to anybody who is a user, information that is posted especially about news or crisis events needs to be validated for credibility. A news item receives numerous responses in a very short span of time and has the potential to create panic related to something that might be totally false. It therefore becomes imperative to have some kind of validity mechanism that can accurately measure the authenticity of a news tweet. This is currently a rich area for research and researchers are contributing actively to the existing body of knowledge on a regular basis.

## 2. RELATED WORK

Twitter has a the potential to spread rumors that could be potentially disastrous and could create a state of panic within a short span of time [5] for example in November 2010 the

twitter account of Indonesian presidential advisor was hacked and then used to post a false tsunami warning similarly in January 2011 rumors regarding shooting in Oxford circus in London were spread rapidly via twitter. Judging from these incidents it becomes absolutely necessary to develop validation mechanisms that could aid in differentiating between true and false information. Following is a brief account of some of the work that has already been done to validate news events credibility.

Kwak et. al. [1] discovered that 85% of discussion topics on twitter are related to news. Zhao et al compared the news tweets of twitter to New York Times using unsupervised topic modeling and showed that twitter users are more active in spreading news of important world events than keeping track of news through television or newspapers. It has been highlighted by various researchers that valuable information can be mined from twitter activity and data in a crisis situation. One of the studies showed that tweets that were initiated on rumor basis had more questions than facts whereas news spreading tweets had factual information that was mostly correct. Another study showed that social networks that belonged to a specific region could provide significant localized information about the event. A study on analysis of twitter data regarding the 2008 Mumbai attacks revealed how the readily available online information aided the terrorist in making crucial decisions. Gupta et al conducted analysis of 14 high impact news events to establish their credibility. The methodology involved querying Trends API after every three hours for trending topics. Over 35 million tweets by more than six million users were collected over a period of 18 days. For the credibility assessment they first took responses like “seems credible” and “seems incredible” from human annotators. The tweets were then fed to the developed system that assessed tweets on the basis of same parameters defined for human responses and then classified them on a credibility ranking. It was statistically validated that the application rating was 75% correct.

Sharma et al [2] developed a “who is who” service for twitter that could accurately infer user attributes and consequently could be effectively used to identify topical experts on twitter. The inference of attributes was mainly done using the Lists feature introduced by twitter for grouping together selective twitter users, thus enabling all the tweets posted by the group to be viewed in the List timeline

Diakopoulous et al [3] designed an information tool for journalists called SRSR (Seriously Rapid Source Review). The application had the ability to search and assess users who could be classified as authentic information source for breaking news. The mechanism for assessment involved various information cues collected from practicing journalist and proved to be quite helpful in the validation process.

Gupta et al [4] used classification techniques for measuring the credibility of tweets. The technique used was to filter out events that had less than a predefined number of tweets. The remaining tweets were assessed on the basis of the comparison of most frequently used tweet words against new

headlines scraped from top ten news websites. Also Google news archive was used for the same purpose.

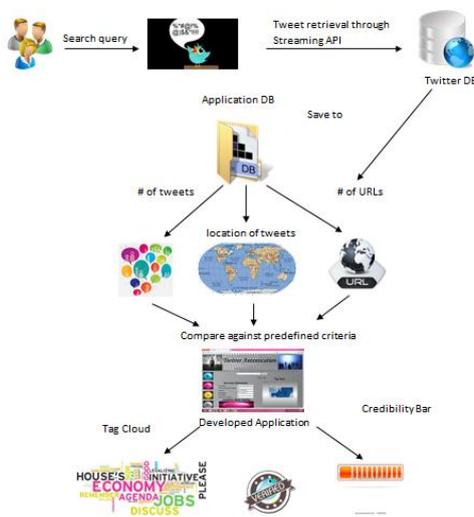
Castillo et al [5] performed automatic credibility analysis by creating a feature list classified as message based, user based, topic based and propagation based. A supervised classifier and J48 learning algorithm were used for the establishment of credibility. Their study also revealed that news related tweets inclined to have more URLs and deeper propagation trees and also that news tweets are generally posted by user who have been actively posting tweets previously and had many followers.

Morris et al [6] proposed assessment of credibility by assigning weights to different features in the related to the users as well as the tweets posted by them. For example features like “is a retweet from someone you trust” and “verified author topic expertise” were assigned more weights leading to higher credibility whereas other features like “non-standard grammar/punctuation” and “default user image” were assigned lower weights. On the whole this study gauges credibility based on obscure information buried in the interface.

### 3. PROPOSED SYSTEM

The system proposed for the credibility assessment of a potential news tweet is heavily inspired by researches already discussed in the previous section. Developed system validates credibility based on following features:

- 1) Number of re-tweets received in a specified interval of time
- 2) The geographical location of the users tweeting and the event are identical
- 3) The tweets contain reference links or URLs for more information regarding the event.



**Figure 1: Description of the developed System**

The first feature is considered authentic based on the reasoning that if a tweet is receiving numerous responses in a very short span of time than that makes it important enough to be treated as a news item. Second feature can help validation in the sense that if most of the tweets are being received from the same geographical region to which the event belongs then there must be a certain degree of truth in the information. The third feature authenticates the tweet content on the basis of sources indicated as references. If the tweets also contain references to other web sources regarding the event then that can also be treated as a validation parameter.

The graphical User Interface performs keyword based search and can be divided into three sections:

- a) Input fields that take search string and date for retrieving tweets in a database. Output fields that return number of tweets in unit time, location of user and event and number of links found in the tweets retrieved.
- b) Graphical representation of credibility assessment result.
- c) Tag Cloud of words used in the tweets retrieved.

Following is the list of features that were taken into account for credibility assessment.

**Table 1. Descriptive statistics**

S No	Field
1	Search Query
2	Total unique users
3	Total Tweets
4	Number of Reference links
5	Number of single tweets
6	Number of retweets
7	Location of event
8	Location of users tweeting



**Figure 2: Information retrieved from tweets database**

The application will first retrieve tweets based on the keyword and date specified by the user. If the user only specifies input string then current date would be used as default value. The tweets will be retrieved continuously unless it is stopped by the user or the application is terminated. The next set of tweets will be retrieved from the time as specified by the last tweet in the retrieved set.

The application will be dynamically retrieving tweets and storing them in the database and the assessment bar and tag cloud will also change dynamically every time a new dataset is retrieved.

This approach is actually a hybrid of methodologies already proposed and implemented and tends to establish authentication of an event using fairly simple but significant features.

### 3.1 Credibility Assessment Criteria

For credibility analysis we trained a supervised classifier using J48 decision tree. Then provided the test data for prediction. The classifier achieves around 80% of accuracy. The data was tested on 100 records and results were found to be accurately rated. Following is test result of the developed application

### 3.2 TwittCounter ANDROID APPLICATION

A simple android application was also developed and uploaded on Google play for testing how accurately the designed application is predicting credibility assessment and also to establish the user response the application might gain.

**Figure 2 : TwittCounter Interface**

Event name	From	To	# of Tweets	Actual value	Resulted value
Barack obama	Aug 24 2013	Aug 30 2013	25	100%	55%
obama	Sep 01 2013	Aug 31 2013	101	100%	99%
junaid jamshaid	Aug 30 2013	Sep 05 2013	4	40%	25%
imran khan	Sep 04 2013	Sep 06 2013	672	100%	99%
mariam khan	Sep 04 2013	Sep 05 2013	4	30%	20%
obama barack	Sep 01 2013	Sep 01 2013	96	70%	65%
obama	Aug 31 2013	Sep 01 2013	101	85%	83%

## 4. CONCLUSION

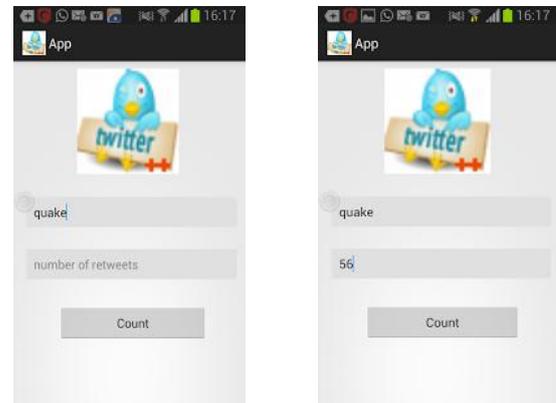
Twitter is no longer used as a medium only for conversing or sharing views but has also become one of the most widely used news disseminating mediums. The tweets posted on twitter are not always true or trustworthy. It therefore becomes essential that effective mechanisms to validate the content posted are in place. The proposed system provides basic credibility checking and can be further enhanced by introducing more features for detailed assessment. Results produced show considerable degree of success in credibility prediction.

The heading of subsections should be in Times New Roman 12-point bold with only the initial letters capitalized. (Note: For subsections and subsubsections, a word like *the* or *a* is not capitalized unless it is the first word of the header.)

## 4. REFERENCES

[1] Aditi Gupta, Ponnurangam Kumaraguru . Credibility Raining of Tweets during High Impact Events. In ACM 2012.  
 [2] Narveen Sharma, Saptarishi Ghosh Fabricio Benevenuto, Niloy Ganguly, Krishna P. Gummadi. Inferring Who-is-Who in the Twitter Social Network. In ACM 2012.

So far the application has been downloaded by more then 300 users and is receiving above average reviews for predicted results.



[3] Nicholas Diakopoulos, Munmun De Choudhury, Mor Naaman. Finding and Assessing Social Media Information Sources in the Context of Journalism. In ACM 2012  
 [4] Maish Gupta, Pexiang Zhao, Jiawei Han. Evaluating Event Credibility on Twitter. In Proc. SDM, 2012, pp.153-164.  
 [5] Carlos Castrillo, Marcelo Mendoza, Barbara Poblete. Information Credibility on Twitter. In ACM 2011  
 [6] Meredith Ringel Morris, Scott Counts, Asta Roseway, Aaron Hoff, Julia Schwarz. Tweeting is Believing? Understanding Microblog Credibility Perceptions. In ACM 2012.  
 [7] Adam Marcus, Michael S. Bernstein, Osama Badar. Processing and Visualizing the Data in Tweets. In SIGMOD Record, December 2011  
 [8] Chunliang Lu, Wai Lam, Yingxiao Zhang. Twitter User Modeling and Tweets Recommendation Based on Wikipedia Concept Graph. In Association for the Advancement of Artificial Intelligence 2012  
 [9] Fabian Abel Qi Gao, Geert-Jan Houben, Ke Tao. Semantic Enrichment of Twitter Posts for user Profile Construction on the Social Web. In In Proceedings of 8th Extended Semantic Web Conference (ESWC2011), Heraklion, Crete, Greece, May 2011.  
 [10] K. R. Canini, B. Suh, and P. L. Pirolli. Finding credible information sources in social networks based on content and social structure. In SocialCom, 2011.  
 [11] A. Gupta and P. Kumaraguru. Twitter explodes with activity in mumbai blasts! a lifeline or an unmonitored daemon in the lurking? In IIIT, Delhi, Technical report, 2011.  
 [12] Oh, M. Agrawal, and H. R. Rao. Information control and terrorism: Tracking the mumbai terrorist attack through twitter. Information Systems Frontiers, 2011.