Analyzing Information on Twitter to Assess Credibility

V. V. S Vineela[1] and Anita Kumari Singh [2]

[1]M.Tech, Department of Computer Science and System Engineering, Andhra University College of Engineering, Visakhapatnam, India

[2]Research Scholar, Department of Computer Science and System Engineering, Andhra University College of Engineering, Visakhapatnam, India

[1](E-mail: vvsvineela445@gmail.com)

[2](E-mail: anitasinghani@gmail.com)

Abstract: *Analyzing the information credibility on social media such as twitter has become an interesting topic among researchers in the fields of both computer and social-sciences. Twitter has become a major platform for rapid dissemination of user-provided content on real time. As such, a large portion of the information it contains is not particularly relevant to many users and in reality it is perceived as unwanted 'noise' by many. There has been increasing interest in predicting whether tweets are relevant, newsworthy or credible using a variety of models and methods. This is critical how people utilize social media changes as time progresses, and the topics discussed vary. In addition, we are constantly gaining new twitter data every day. Thus, it is important to be able to have a set of features that can perform well across many different topics. The proposed system consists of four integrated components: a reputation based model, a feature ranking algorithm, a credibility classification engine and a user expertise component. The components operate in an algorithmic form to analyze and assessing the credibility of Twitter tweets and Users. We tested the performance of our system using 5000 different Twitter tweets with unique users. We applied 10-fold cross validation on the machine learning algorithms.*

Index terms: *Reputation, feature-ranking, credibility, user expertise, Twitter*

1 Introduction

Studying information credibility on twitter has become a popular topic. Can one determine whether information found on twitter is credible or not? Which specific features are most relevant to identifying credibility of content? In parallel with these studies are those that try to determine whether features obtained from twitter activity can be used to determine who the experts are, or who the influential people are. When studying credibility it is important to consider both the type of data that was used in analysis as well as the methods used to generate ground truth. The types of features available for study are large and they can be studied using linear regression, logistic regression etc.

To motivate this problem, let us consider the problem of information credibility. Suppose a twitter study show that four most important factors that determine information credibility are: presence of URLs, the number of followers the user has, the numbers of friends the user has and the number of positive and negative sentiments expressed. This raises a number of questions.

Suppose we would like to study credibility in a specific scenario, i.e. an earthquake in turkey. The user population in our study may be different than those present in this scenario. Some features may be culturally dependent: in this case, everyone may have emotional responses regardless of whether the content is credible or not. The presence of URLs may not be very useful at all as the information is being generated very quickly to even post on a site. In fact, most long term power users may be far removed from the actual event and the world becomes here say by the time it reaches them. Overall, these features may not be useful at all to determine credibility in this scenario.

Currently researchers have developed various methodologies in studies on information credibility. Some of them consider the problem to be one of classification that should be solved in an automated fashion using machine learning or graph-based algorithms. Others view it as cognitive problem requiring human-centric verification. Other researchers have ventured to devise algorithms for assessing credibility, while others have studied the visualization of credibility scores using radar graphs and comparisons between systems such as Fluo and TopicNets. Such systems include TweetCred and Twitter-Trails.

- Our proposed system consists of four main integrated components in reaching a precise information credibility judgement. The components are a feature ranking algorithm, a reputation based model, a credibility assessment classifiers engine and a user expertise model.
- Using feature ranking algorithm, we sought to automatically rank which features are important from the extracted features.
- Using the reputation based technique, we automatically rank users and tweets based on their relevance and expertise.
- Credibility classification technique is used to check whether the tweet declared as credible or not is correct or not. Our, priority being to minimize false positives, we might choose to optimize our model with respect to recall or sensitivity.
- We validated our system by applying tenfold cross-validation with four machine learning algorithms on our dataset.

2 Methodology

2.1 Data Collecting

Data can be collected through two different Twitter application interfaces (APIs) either using streaming API or through searching API. The streaming API is used to collect datasets on given events. The search API is used to collect users tweet histories simultaneously.

In order to extract tweets from the twitter platform you will need a twitter application and hence a twitter account. After the successful creation of twitter application you need to install required packages and then establish a secure connection using the keys which are generated while creating the twitter application.

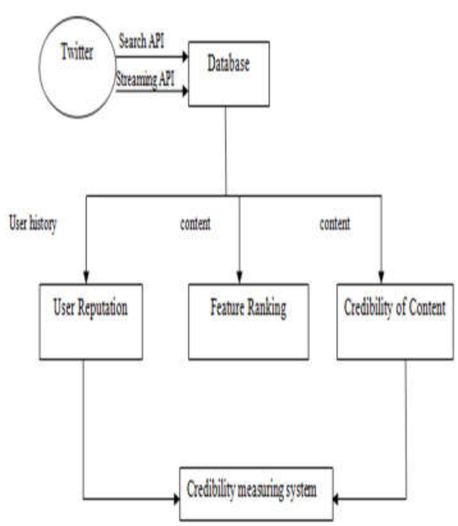


Fig 2.1 Block diagram of credibility analysis system

After the establishment of secure connection now search for the tweets of your required event using searchTwitter() function and store them in required format. After successful collection of the data now pre-process the data for removing stop-words, unwanted characters etc,. The prepared data is then passed as inputs to three techniques to look for signals of credibility.

2.2 User reputation

Calculating user reputation is an important aspect of problem to solve because this case of vision is widely spread, especially on social-networks. To calculate user reputation we use different measures that have an enormous amount of impact on twitter. This can be achieved by measuring reputation through how popular a user is and how much sentiment expressed in his/her tweets.

Sentiment Calculation sentiment scoring technique mainly aims in determining whether a tweet was ‘positive’, ‘negative’ or ‘neutral’. This can be achieved by using predefined positive and negative words. By using the positive and negative words we first count how many positive and negative words are present in the tweet. By taking subtraction of sum of positive words and sum of negative words we get a sentiment score of a tweet.

Sentiment score=sum (positive words)-sum (negative tweets)

After calculating sentiment score then we can define the tweet as positive or negative or neutral basing on the score.

User Reputation user reputation is the most important aspect of the credibility analysis. User reputation can be achieved by extracting the history of the user. The number of followers and friends that one has on twitter impacts the popularity and trust-worthiness of a user. Twitter and Face-book has somewhat different mechanics for collecting social connections. A face-book friend represents a bi-directional relationship whereas on twitter one may choose whom to follow and not. And the person being followed may choose to remove a follower.

The ratio of this number of followers that a twitter user has and how many others they follow has a potential impact on credibility of a user. If the resultant is greater than some value i.e. greater than or equal to five then we can declare that the user who posted that content is a credible user.

2.3 Feature Ranking Algorithm

The extracted features are divided into three levels they are Tweet-level, User level and hybrid level. The tweet level features include some of the characteristics such as length of a message, number of re-tweets, hash-tags, user mentions as well as URLs and number of static and animated emotions. User level features include number of followers, number of friends, age, gender, as well as replies of user's tweets. Hybrid level features include all the tweet level and user level features.

The extracted features should be weighted before measuring the assessment of a given tweet, because of the influence of the features on final judgement of credibility. We deduced that the number of friends a user has is the most influencing feature, followed by number of followers, re-tweets and user mentions.

The least influential factors are considered to be location, time zone and the number of favourites. This overall feature importance can be obtained by using logistic regression technique. In this method first we constructed a model and after that we apply the logistic regression technique to the model.

2.4 Credibility classification engine

The credibility classification engine is used to calculate the credibility assessment. This can be achieved by applying supervised classification to the collection to guarantee high recall. The main hypothesis is we can automatically estimate the credibility score of the information collected through twitter.

We trained a supervised classifier to determine the credibility of each tweet. Here we used a well known classification algorithm Naive Bayesian classification algorithm. It is a model that assigns class labels to problem instances, represented as vector of the feature values, where the class labels are drawn from finite set.

To apply the classification algorithm to our dataset first we need to calculate document term matrix to our dataset. After calculating the document term matrix we need to divide the dataset into train dataset and test dataset. Now train the train dataset using naive Bayesian classifier. After training the classifier now predict the calculations using test dataset. Build a confusion matrix for the predicted values. Finally calculate the overall accuracy of the observations.

3 Results

A. The following screen shows extracting the tweets from using search API.

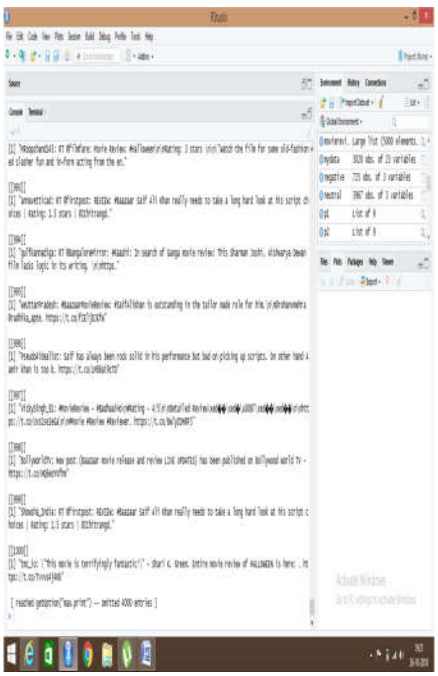


Fig 3.1 Extracting tweets

B. Measuring the sentiment score of each tweet has been shown in the following screen.

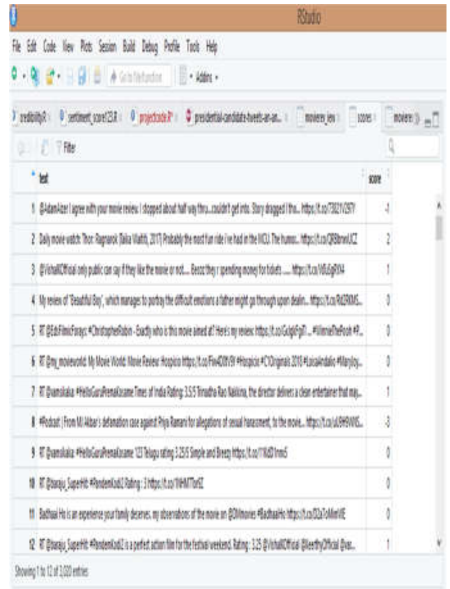
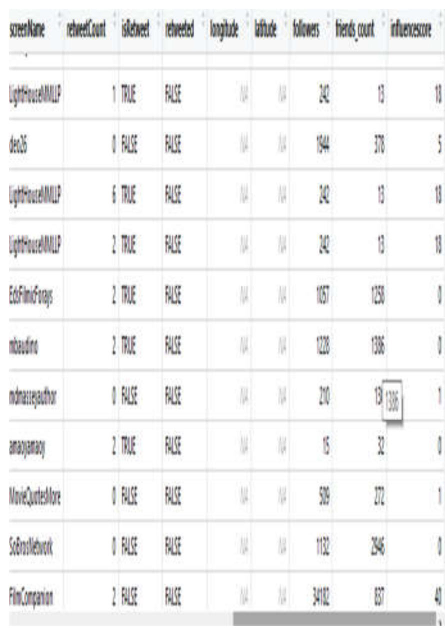


Fig 3.2 Sentiment score of each tweet

C. The influence score of each user is shown in the following screen



screenName	retweetCount	isRetweet	retweeted	longitude	latitude	followers	friends_count	influencescore
LightHouseMMLP	1	TRUE	FALSE	NA	NA	242	13	18
deo26	0	FALSE	FALSE	NA	NA	1944	378	5
LightHouseMMLP	6	TRUE	FALSE	NA	NA	242	13	18
LightHouseMMLP	2	TRUE	FALSE	NA	NA	242	13	18
EdsFilmicForays	2	TRUE	FALSE	NA	NA	1057	1258	0
mbaudino	2	TRUE	FALSE	NA	NA	1228	1386	0
mdmasseyauthor	0	FALSE	FALSE	NA	NA	210	13	1
amaoyamaoy	2	TRUE	FALSE	NA	NA	15	32	0
MovieQuotesMore	0	FALSE	FALSE	NA	NA	509	272	1
SciBrosNetwork	0	FALSE	FALSE	NA	NA	1132	2946	0
FilmCompanion	2	FALSE	FALSE	NA	NA	34182	837	40

Fig 3.3 influence score of each user

D. For calculating accuracy we applied naive Bayesian algorithm to our model. The following screen shows the accuracy of the model.

Fig: Accuracy of the model

4 Conclusion

This paper presents the problem of assessing the information credibility on twitter. This problem of information credibility has come under survey, especially in social networks. We used user history and the sentiments of the tweets to solve the problem of assessing information credibility. The effectiveness of the system is tested using the tenfold cross validation over the machine learning algorithms.

In the near future we will try to analyze the problem of credibility using time sensitive and location-based approaches which gives more reliable and trustworthiness results.

References

[1] Gupta, P. Kumaraguru, C. Castillo, and P. Meier, "Tweetcred: Real-time credibility assessment of content on twitter," in Social Informatics, ed: Springer, 2014, pp. 228-243

[2] M. R. Morris, S. Counts, A. Roseway, A. Hoff, and J. Schwarz, "Tweeting is believing?: understanding microblog credibility perceptions," in Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, 2012, pp. 441 450.

[3] M. Al-Qurishi, R. Aldrees, M. AlRubaian, M. Al-Rakhami, S. M. M. Rahman, and A. Alamri, "A new model for classifying social media users according to their behaviors," in Web Applications and Networking (WSWAN), 2015 2nd World Symposium on, 2015, pp. 15.

[4] A.AlMansour, L. Brankovic, and C. S. Iliopoulos, "Evaluation of credibility assessment for microblogging: models and future directions," in Proceedings of the 14th International Conference on Knowledge Technologies and Data-driven Business, 2014, p. 32.

[5] M. Abdul-Mageed, M. Diab, and S. Kubler, "SAMAR: Subjectivity and sentiment analysis for Arabic social

media," Computer Speech & Language, vol. 28, pp. 20-37, 2014.

[6] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," presented at the Proceedings of the 20th international conference on World wide web, Hyderabad, India, 2011.

[7] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida. Detecting spammers on Twitter. In CEAS, 2010.

[8] A.L. Hughes and L. Palen. Twitter adoption and use in crisis twitter adoption and use in mass convergence and emergency events. In ISCRAM, 2010.

[9] Metaxas, Panagiotis Takas, Samantha Finn, and Eni Mustafaraj. "Using TwitterTrails. com to Investigate Rumor Propagation." Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing. ACM, 2015.