**A Credibility Analysis System for Assessing Information on Twitter**

**ABSTRACT:**

Information credibility on Twitter has been a topic of interest among researchers in the fields of both computer and social sciences, primarily because of the recent growth of this platform as a tool for information dissemination. Twitter has made it increasingly possible to offer near-real-time transfer of information in a very cost-effective manner. It is now being used as a source of news among a wide array of users around the globe. The beauty of this platform is that it delivers timely content in a tailored manner that makes it possible for users to obtain news regarding their topics of interest. Consequently, the development of techniques that can verify information obtained from Twitter has become a challenging and necessary task. In this paper, we propose a new credibility analysis system for assessing information credibility on Twitter to prevent the proliferation of fake or malicious information. The proposed system consists of four integrated components: a reputation-based component, a credibility classifier engine, a user experience component, and a feature-ranking algorithm. The components operate together in an algorithmic form to analyze and assess the credibility of Twitter tweets and users. We tested the performance of our system on two different datasets from 489,330 unique Twitter accounts. We applied 10-fold cross-validation over four machine learning algorithms. The results reveal that a significant balance between recall and precision was achieved for the tested dataset.

**INTRODUCTION**

Online social networks, such as Twitter, have grown highly popular in the 21st century, as the numbers of users who are using them on daily basis attest. Information dissemination through these platforms is their most attractive feature, as it is known to be speedy and cost effective. The fact that users are allowed to express themselves with little to no control is also another very attractive aspect of these platforms [1]. As users are afforded the freedom to publish content with no supervision, the problem of information credibility on social networks has also risen in recent years. Crafty users of these platforms can spread in-formation maliciously for reasons that may not be compatible with the good of society. Users are becoming wary that rumors that are spread through online social networks can have detrimental effects. Research on information credibility is thus the best solution to the problem of how to assess the credibility of information

**PROBLEM STATEMENT**

* Currently, researchers have employed 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 a cognitive problem requiring human-centric verification.
* Some authors have looked at how various aspects of social media, such as the effect of the name value and user connectedness, influence users’ judgments concerning credibility
* Other researchers have ventured to devise algorithms for assessing credibility, while others have studied the visualization of credibility scores using such means as radar graphs and comparisons between systems such as Fluo and TopicNets.
* Some researchers have gone so far as to create systems to assess credibility automatically in real time. Such systems include TweetCred and Twitter-Trails.

**DISADVANTAGES OF PROBLEM STATEMENT:**

* The complexity of social networks and the web creates difficulty in identifying resources for use in studying and assessing credibility.
* OSNs by their very nature evolve dynamically over time and become very large in size, with various structures that make it difficult to obtain the information needed to discern the credibility of users.
* The credibility of a user is influenced continuously by various factors, such as changes in the social topography, other users’ behavior, preferences, and context.
* Malicious activities can evade existing spam filters through various means. For example, in Twitter, malicious users can purchase followers or use tools to automatically generate fake ac-counts and post tweets with the same meaning but different words.
* The process of evaluating solutions has also been a problem in terms of resources, given that most researchers are limited in terms of the extent to which they can test their work.

**LITERATURE SURVEY**

In this section, we briefly outline some of the research work done to assess, characterize, analyze, and compute trust and credibility of content in online social media.

**Credibility Assessment. Castillo et al. [4]** showed that automated classification techniques can be used to detect news topics from conversational topics and assess their credibility based on various Twitter features. They achieved a precision and recall of 70-80% using a decision-tree based algorithm.

**Gupta and Kumaraguru [7]** in their work on analyzing tweets posted during the terrorist bomb blasts in Mumbai (India, 2011), showed that the majority of sources of information are unknown and have low Twitter reputation (small number of followers). The authors in a follow up study applied machine learning algorithms (SVM-rank) and information retrieval techniques (relevance feedback) to assess credibility of content on Twitter [8], finding that only 17% of the total tweets posted about the event contained situational awareness information that was credible.

**Another, similar work was done by Xia et al. [19]** on tweets generated during the England riots of 2011. They used a supervised method based on a Bayesian Network to predict the credibility of tweets in emergency situations.

**O’Donovan et al. [15]** focused their work on finding indicators of credibility during different situations (8 separate event tweets were considered). Their results showed that the best indicators of credibility were URLs, mentions, retweets and tweet length.

**Credibility perceptions. Morris et al. [14]** conducted a survey to understand users’ perceptions regarding credibility of content on Twitter. They found that the prominent features based on which users judge credibility are features visible at a glance, for example, the username and picture of a user.

**Yang et al. [21]** analyzed credibility perceptions of users on two micro-blogging websites: Twitter in the USA and Weibo in China. They found that location and network overlap features had the most influence in determining the credibility perceptions of users.

**Aditi Gupta et al. Credibility of users. Canini et al. [3]** analyzed the usage of automated ranking strategies to measure credibility of sources of information on Twitter for any given topic. The authors define a credible information source as one which has trust and domain expertise associated with it.

**Ghosh et al. [6]** identified topic based experts on Twitter using features obtained from user-created list, relying on the wisdom of Twitter’s crowds.

**System. Ratkiewicz et al. [17]** introduced Truthy,4 a system to study information diffusion on Twitter and compute a trustworthiness score for a public stream of micro-blogging updates related to an event. Their focus is to detect political smears, astroturfing, and other forms of politically-motivated disinformation campaigns.

To the best our knowledge, the work presented in this paper is the first research work that describes the creation and deployment of a practical system for credibility on Twitter, including the evaluation of such system with real users

**OBJECTIVE OF THE PROPOSED RESEARCH:**

* In this paper, we make a practical approach to automated credibility assessment on Twitter. We describe the process behind the design of an automated classifier for information credibility assessment.
* As an addition, we propose practical implementation of Twitter BOT, a tool which is able to score submitted tweets while working in the native Twitter interface

**METHODOLOGY OF THE PROPOSED RESEARCH:**

A system was developed to measure the credibility of news content published in Twitter. The system uses two approaches to assign credibility levels (low, high and average) to each tweet. The first approach is based on the similarity between Twitter posts (tweets) and authentic (i.e. verified) news sources. The second approach is based on the similarity with verified news sources in addition to a set of proposed features.

**PROPOSED SYSTEM:**

* We propose a hybrid approach to credibility analysis that can be used to identify implausible content on Twitter and prevent the proliferation of fake or malicious information.
* We propose a novel credibility assessment system that maintains complete entity-awareness (tweet, user) in reaching a precise information credibility judgment.
* This model comprises four integrated components, namely, a reputation- based model, a feature ranking algorithm, a credibility assessment classifiers engine, and a user expertise model.
* All of these components operate in an algorithmic form to analyze and assess the credibility of the tweets on Twitter.
* In our system, an observation is a tweet, and the positive class is credible. In this case, a highly sensitive classifier is more acceptable than precision, because non-credible tweets, if classified as credible, might spread misinformation that goes viral and cause chaos in terms of politics or an emergency. Thus, our priority being to minimize false positives, we might choose to optimize our model with respect to recall or sensitivity.

**EXPECTED OUTCOME OF THE PROPOSED RESEARCH:**

This paper contributes to the research on web credibility. It is believed to be the first which provides a proposed system to evaluate the credibility of Twitter news content automatically.

**ADVANTAGES OF PROPOSED SYSTEM:**

* Using the reputation-based technique, we sought to automatically rank users based on their relevance and expertise on given topics.
* We enhanced our classifier by weighing each feature according to its relative importance. This weighting methodology implements a pairwise comparison that produces a priority vector that ranks the instance’s features ac-cording to their relative importance with respect to the user need as well as the topic.
* We validated our system by applying tenfold cross validation and on different datasets of Twitter content. Our results show that the system that employed a reputation-based filter approach provide a significant and accurate credibility assessment.

**SYSTEM REQUIREMENTS:**

**HARDWARE REQUIREMENTS:**

* System : Pentium Dual Core.
* Hard Disk : 120 GB.
* Monitor : 15’’ LED
* Input Devices : Keyboard, Mouse
* Ram : 1 GB

**SOFTWARE REQUIREMENTS:**

* Operating system : Windows 7.
* Coding Language : JAVA/J2EE
* Tool : Netbeans 7.2.1/Eclipse IDE
* Database : MYSQL

**SUMMARY:**

Owing to the large amount of information available on Twitter (a micro blogging service) that is not necessarily true or believable, credibility of news published in such an electronic channel has become an important area for investigation in the field of web credibility. This paper aims to address this issue.

**REFERENCES**

[1] 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. 1- 5

[2] M. AlRubaian, M. Al-Qurishi, M. Al-Rakhami, S. M. M. Rahman, and A. Alamri, "A Multi-stage Credibility Analysis Model for Microblogs," presented at the Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015, Paris, France, 2015

[3] A. 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.

[4] Majed AlRubaian, Muhammad Al-Qurishi, Sk Md Mizanur Rahman, and A. Alamri, "A Novel Prevention Mechanism for Sybil Attack in Online Social Network," presented at the WSWAN'2015, 2015

[5] 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.