**Spammer Detection and Fake User Identification on Social Networks**

**Abstract:**

Social networking sites engage millions of users around the world. The users’ interactions with these social sites, such as Twitter and Facebook have a tremendous impact and occasionally undesirable repercussions for daily life. The prominent social networking sites have turned into a target platform for the spammers to disperse a huge amount of irrelevant and deleterious information. Twitter, for example, has become one of the most extravagantly used platforms of all times and therefore allows an unreasonable amount of spam. Fake users send undesired tweets to users to promote services or websites that not only affect legitimate users but also disrupt resource consumption. Moreover, the possibility of expanding invalid information to users through fake identities has increased that results in the unrolling of harmful content. Recently, the detection of spammers and identification of fake users on Twitter has become a common area of research in contemporary online social Networks (OSNs). In this paper, we perform a review of techniques used for detecting spammers on Twitter. Moreover, a taxonomy of the Twitter spam detection approaches is presented that classifies the techniques based on their ability to detect: (i) fake content, (ii) spam based on URL, (iii) spam in trending topics, and (iv) fake users. The presented techniques are also compared based on various features, such as user features, content features, graph features, structure features, and time features. We are hopeful that the presented study will be a useful resource for researchers to find the highlights of recent developments in Twitter spam detection on a single platform.

**Existing system:**

In existing system is to identify different approaches of spam detection on Twitter and to present a taxonomy by classifying these approaches into several categories. For classification, we have identified four means of reporting spammers that can be helpful in identifying fake identities of users. Spammers can be identified based on: (I) fake content, (ii) URL based spam detection, (iii) detecting spam in trending topics, and (iv) fake user identification. Provides a comparison of existing techniques and helps users to recognize the significance and effectiveness of the proposed methodologies in addition to providing a comparison of their goals and results. Compares different features that are used for identifying spam on Twitter. We anticipate that this survey will help readers find diverse information on spammer detection techniques at a single point.

**Disadvantages:**

* Efficient and effective approaches
* More time

**Proposed System:**

The proposed approach utilizes the information contained in the tweets when a spam or malware is recognized by the users or the report of security has been released by the certified authorities. The proposed alerting system comprises of the following components: (i) real time data extraction of both tweets and users, (ii) filtering system based on a preprocessing schedule and on Naïve Bayes algorithm to discard the tweets containing inaccurate information, (iii) data analysis for spammer detection where the detection windows a are rigorously abolished according to the Sigmoid function or when the window size reaches the maximum, (iv) alert subsystem that is used when the event is established, the system groups up the tweets that are relevant to the same topic where tweets are distinguished with the cluster barycenter and the one that is nearest to the cluster center is chosen as the representative of the whole system cluster, and (v) feedback analysis. The approach is claimed to be efficient and effective for the detection of some invasive and admirable malignant activities in circulation.

**Advantages:**

* Less time
* More efficient

**S/W System Configuration**

**Software Requirements**

* Operating System : Windows 8
* Application Server : Tomcat 7.0
* Front End : HTML, JSP,CSS
* Scripts : JavaScript.
* Backend Language : Java
* Database : MySQL 6.0
* IDE : Eclipse(2019-3)

**Hardware Requirements**

# Processor - Intel i3

* RAM - 4GB
* Hard Disk - 500 GB