International Journal of Engineering Research and Science & Technology

IJERST

ISSN : 2319-5991

www.ijerst.com

Email: editor@ijerst.com or editor.ijerst@gmail.com

Int. J. Eng. Res. & Sci. & Tech. 2025

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

ENHANCING SOCIAL NETWORK SECURITY: SPAMMER DETECTION AND FAKE USER IDENTIFICATION

K. Ram Mohan, P. Swetha, A. Rajeshwari, B. Vanaja

[1]Assistant Professor, Department of Computer Science and Engineering,

[1,]St. Martin's Engineering College, Secunderabad, Telangana, India

ABSTRACT

This project aims to delineate a methodology for detecting spam tweets and fraudulent user accounts on the social network Twitter. We are utilizing a Twitter dataset and employing four distinct techniques: Fake Content Detection, Spam URL Detection, Spam Trending Topic Identification, and Fake User Identification. By employing the aforementioned four techniques, we can ascertain whether a tweet is normal or spam. Subsequently, we will utilize the Random Forest data mining algorithm to train the dataset for the classification of spam versus non-spam tweets, as well as fake versus non-fake accounts. We employ various data mining techniques to classify tweets as spam or non-spam, specifically utilizing the Random Forest classifier in this instance. There is a demand to address and regulate individuals who utilize online social networks solely for advertising, thereby spamming others' accounts. The recent identification of spam on social networking platforms has garnered the interest of researchers. Spam detection poses a significant challenge in safeguarding the security of social networks. Recognizing spam on OSN sites is crucial to protect users from diverse malicious attacks and to safeguard their security and privacy.

Keywords: Spam tweets, fraudulent user accounts, trending topics, data set, and random forest data mining algorithm.

1. INTRODUCTION

It has become quite unpretentious to obtain any kind of information from any source across the world by using the Internet. The increased demand of social sites permits users to collect abundant amount of information and data about users. Huge volumes of data available on these sites also draw the attention of fake users. Twitter has rapidly become an online source for acquiring real-time information about users. Twitter is an Online Social Network (OSN) where users can share anything and everything, such as news, opinions, and even their moods. Several arguments can be held over different topics, such as politics, current affairs, and important events. When a user tweets something, it is instantly conveyed to his/her followers, allowing them to outspread the received information at a much broader level . With the evolution of OSNs, the need to study and analyze users' behaviors in online social platforms has intensity. Many people who do not have much information regarding the OSNs can easily be tricked by the fraudsters. There is also a demand to combat and place a control on the people who use OSNs only for advertisements and thus spam other people's accounts. Recently, the detection of spam in social networking sites attracted the attention of researchers. Spam detection is a difcult task in maintaining the security of social networks.

It is essential to recognize spams in the OSN sites to save users from various kinds of malicious attacks and to preserve their security and privacy. These hazardous maneuvers adopted by spammers cause massive destruction of the community in the real world. Twitter spammers have various objectives, such as spreading invalid information, fake news, rumors, and spontaneous messages. Spammers achieve their malicious

objectives through advertisements and several other means where they support different mailing lists and subsequently dispatch spam messages randomly to broadcast their interests. These activities cause disturbance to the original users who are known as non-spammers. In addition, it also Decreases the repute of the OSN platforms. Therefore, it is essential to design a scheme to spot spammers so that corrective efforts can be taken to counter their malicious activities. Several research works have been carried out in the domain of Twitter spam detection. To encompass the existing state-of the- art, a few surveys have also been carried out on fake user identification from Twitter. Ting min et al. provides a survey of new methods and techniques to identify Twitter spam detection. The above survey presents a comparative study of the current approaches. On the other hand, the authors in conducted a survey on different behaviors exhibited by spammers on Twitter social network. The study also provides a literature review that recognizes the existence of spammers on Twitter social network. Despite all the existing studies, there is still a gap in the existing literature. Therefore, to bridge the gap, we review state- of-the- art in the spammer detection and fake user identication on Twitter. Moreover, this survey presents taxonomy of the Twitter spam detection approaches and attempts to offer a detailed description of development in the domain.

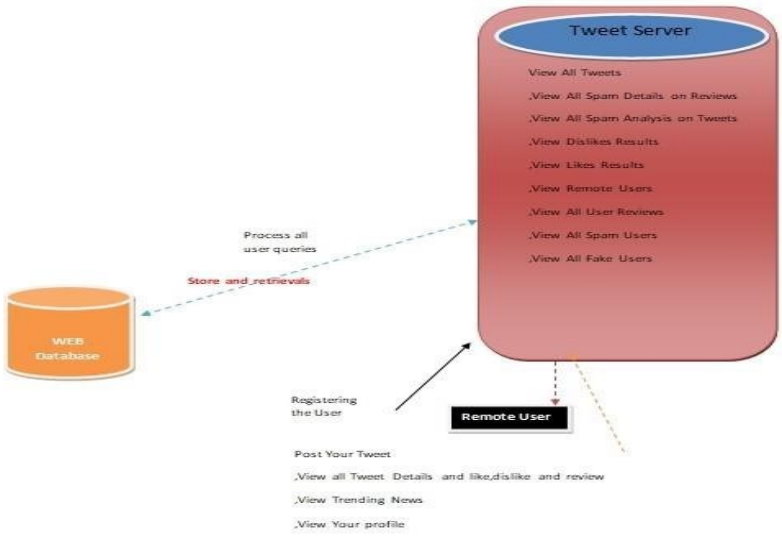

Fig:1 Architecture Diagram

The aim of this paper is to identify different approaches of spam detection on Twitter and to present 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 ,(iii) detecting spam in trending topics, and (iv) fake user identification. Table 1 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. Table 2 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.

2. SPAMMER DETECTION ON TWITTER

URL In this article, we elaborate a classification of spammer detection techniques. Fig. 2 shows the proposed taxonomy for identification of spammers on Twitter. The proposed taxonomy is categorized into four main classes, namely, (i) fake content; (ii) URL based spam detection, (iii) detecting spam in trending topics, and
(iv) fake user identification. Each categoryof identification methods Relies on a specific model, technique, and detection algorithm. The first category (fake content) includes various techniques, such as regression prediction model, malware alerting system, and Lfun scheme approach. In the second category based spam detection),the spammer is identified in URL through different machine learning algorithms. The third category (spam in trending topics) is identified through Naïve Bayes classifier and languagemodel divergence. The last category (fake user identification) is based on detecting fake users through hybrid techniques. Techniques related to each of the spammer identification categories are discussed in the following subsections.

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

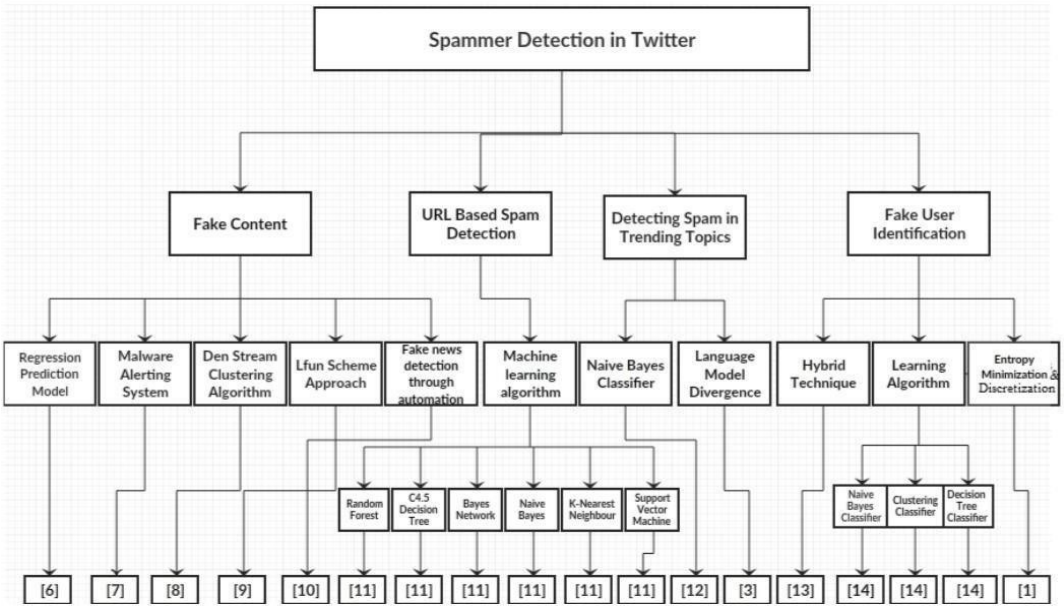

Fig: 2: The proposed taxonomy for identification of spammers on Twitter

Proposed Method	Goal	Data Set	Result
Dirichlet distribution has been used by the statistical framework for identifying spammer in Twitter.	Distinguish between spammer and non-spammer	Real data of Twitter and Instagram	Experimentation carried out on Instagram and Twitter data shows that supervised and unsupervised algorithmic methods deliver meaningful outcomes.
Effective unified weighted for anomalous URL detection	Detection of anomalies behavior in user's interaction	Twitter dataset is used, which contains last 200 tweets of users	Anomalous detection model can be used to analyze effectively the number of URL spammer that is done every day.
Using manual inspection, classification of users as spammer and non-spammer	Detection of spammer on Twitter	Twitter dataset that includes more than 1.9 billion links and tweets around 1.8 billion.	Classification of spammer uses a large set of attributes and is significantly more robust to spammers, which familiarize spamming schemes.
Three types of cascade information, which are created on the basis of spam detection mechanism, have been used, i.e., TSP, SS, and cascade filtering.	Spammers have been classified by using the properties of social networks in the individual social environment.	Real Twitter dataset.	The schemes are scalable because they check users cantered 2-hops social networks instead of examining the whole network.
Design of 18 robust features by holding the time properties explicitly and implicitly.	Answer the question of how to identify spammer only	Crawled and manually annotated dataset	The features extracted are able to recognize both authentic users and spammers accu-

FIG: 3 comparisons between proposed methods for spam detection in twitter

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

3. RESULT ANALYSIS:

In the proposed system, the system elaborates a classification of spammer detection techniques. The system shows the proposed taxonomy for identification of spammers on Twitter. The proposed taxonomy is categorized into four main classes, namely, (i) fake content, (ii) URL based spam detection, (iii) detecting spam in trending topics, and (iv) fake user identification. Each category of identification methods relies on a specific model, technique, and detection algorithm. The first category (fake content) includes various techniques, such as regression prediction model, malware alerting system, and Lfun scheme approach. In the second category (URL based spam detection), the spammer is identified in URL through different machine learning algorithms. The third category (spam in trending topics) is identified through Naïve Bayes classifier and language model divergence. The last category (fake user identification) is based on detecting fake users through hybrid techniques.

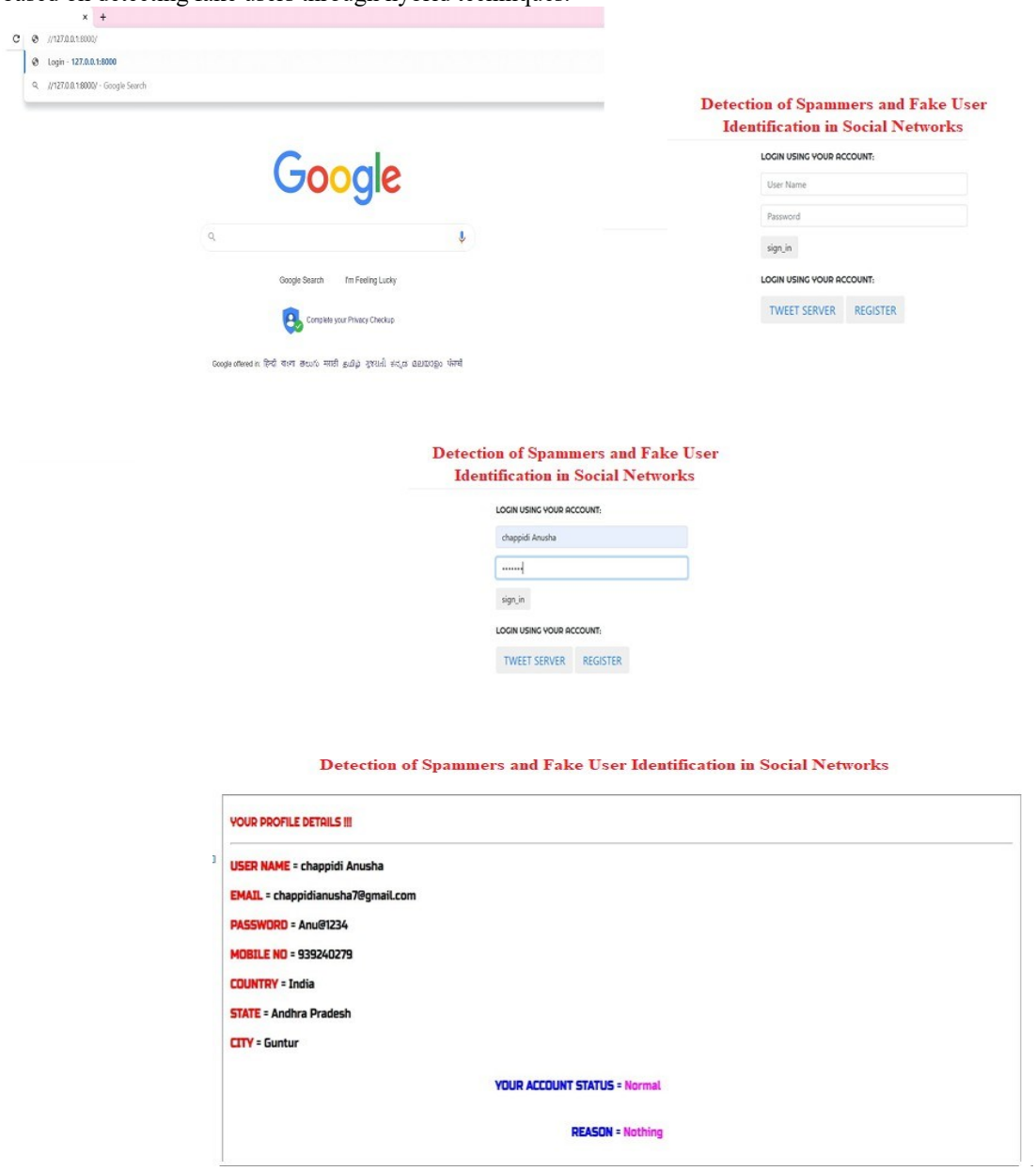

Fig: 4 Few Results of Detection of Spammers and Fake User Identification in Social Networks.

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

4. CONCLUSION

In this paper, we performed a review of techniques used for detecting spammers on Twitter. In addition, we also presented taxonomy of Twitter spam detection approaches and categorized them as fake content detection, URL based spam detection, spam detection in trending topics, and fake user detection techniques. We also compared the presented techniques based on several features, such as user features, content features, graph features, structure features, and time features. Moreover, the techniques were also compared in terms of their specified goals and datasets used. It is anticipated that the presented review will help researchers find the information on state-of-the-art Twitter spam detection techniques in a consolidated form.

Despite the development of efficient and effective approaches for the spam detection and fake user identification on Twitter, there are still certain open areas that require considerable attention by the researchers. The issues are briery highlighted as under: False news identification on social media networks is an issue that needs to be explored because of the serious repercussions of such news at individual as well as collective level. Another associated topic that is worth investigating is the identification of rumor sources on social media. Although a few studies based on statistical methods have already been conducted to detect the sources of rumors, more sophisticated approaches, e.g., social network based approaches, can be applied because of their proven effectiveness

5. REFERENCES

[1] J. Li, M. Ott, C. Cardie and E. Hovy, “Towards a General Rule for Identifying Deceptive Opinion Spam,” in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, MD, USA, vol. 1, no. 11, pp. 1566-1576, November 2014.

[2] Gajjala Buchi Babu, Mutyala Venu Gopal, Vellala Sai Srinivas, V. Krishna Pratap, Efficient Key Generation for Multicast Groups Based on Secret Sharing, (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 1, Issue 4, pp.1702-1707.

[3] Karthik, J. V., & Reddy, B. V. (2014). Authentication of secret information in image stenography. International Journal of Computer Science and Network Security (IJCSNS), 14(6), 58.

[4] Veerendra, B., & Reddy, B. V. (2013). Implementation of Skyline Sweeping Algorithm. International Journal of Computer Science & Engineering Technology (IJCSET) ISSN, 2229-3345.