

MACHINE LEARNING FRAMEWORK FOR DETECTING SPAMMER AND FAKE USERS

Mr. M.Sunil Kumar

Asst.Professor

Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

P.Poojitha

Student

Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

M.Sravani

Student

Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

Pranavi joshi

Student

Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

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 the 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 spamming. Fake users send undesired tweets to users to promote services or websites that not only affect the legitimate users but also disrupt the 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 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

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 difficult 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 reputation 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.

EXISTING SYSTEM

- Tingmin et al. provide 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, S. J. Soman et. al. conducted a survey on different behaviours 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 identification on Twitter

Disadvantages –

- No efficient methods used.
- No real time data is used.
- More complex

PROPOSED SYSTEM

- 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

Advantages • This study includes the comparison of various previous methodologies proposed using different datasets and with different characteristics and accomplishments.

- Tested with real time data.
- The fake content propagation was identified through the metrics that include: (i) social reputation, (ii) global engagement, (iii) topic engagement, (iv) likability, and (v) credibility. After that, the authors utilized regression prediction model to ensure the overall impact of people who spread the fake content at that time and also to predict the fake content growth in future

SYSTEM ENVIRONMENT

Python:

It is an interpreted, high-level and general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python interpreters are available for many operating systems. Python is managed by the non-profit Python Software Foundation. Python features a dynamic types system and automatic memory management. It supports multiple programming paradigms, including object oriented, functional and procedural and has a large and comprehensive standard library. Python is easy to learn yet powerful and versatile scripting language, which makes it attractive for Application Development.

Python IDLE:

IDLE stands for Integrated Development and Learning Environment. The story behind the name IDLE is similar to Python. Guido Van Rossum named Python after the British comedy group Monty Python while the name IDLE was chosen to pay tribute to Eric Idle, who was one of the Monty Python's founding members. IDLE comes bundled with the default implementation of the Python language since the 01.5.2b1 release. It is packaged as an optional part of the Python packaging with many Linux, Windows, and Mac distributions.

DJANGO:

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open source. Django's primary goal is to ease the creation of complex, database-driven websites. Django emphasizes reusability and "pluggability" of components, rapid development, and the principle of don't repeat yourself. Python is used throughout, even for settings files and data models.

IMPLEMENTATION

Service Provider :

In this module, the Admin has to login by using valid user name and password. After login successful he can perform some operations such as View All tweets, Find the type of tweets, View All Registered Users, View All tweet results, Extract spammers and Fake users.

Viewing and Authorizing Users:

In this module, the provider views all users details and authorize them for login permission. User Details such as User Name, Address, Email Id and Mobile Number.

User:

In this module, there are n numbers of users are present. Users should register before performing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user can perform some operations like BROWSE

TWEETS, VIEW ALL TWEETS DETAILS, SEARCH FOR THE TWEETS.

Viewing Profile Details:

details,such as their address, email, mobile number, profile Image.

CONCLUSION

In this project, we performed a review of techniques used for detecting spammers on Twitter. In addition, we also presented a 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.

REFERENCES

- [1] B. Erçahin, Ö. Akta³, D. Kiliç, and C. Akyol, "Twitter fake account detection," in Proc. Int. Conf. Comput. Sci. Eng. (UBMK), Oct. 2017, pp. 388392.
- [2] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on Twitter," in Proc. Collaboration, Electron. Messaging, AntiAbuse Spam Conf. (CEAS), vol. 6, Jul. 2010, p. 12.
- [3] S. Gharge, and M. Chavan, "An integrated approach for malicious tweets detection using NLP," in Proc. Int. Conf. Inventive Commun. Comput. Technol. (ICICCT), Mar. 2017, pp. 435438.
- [4] T. Wu, S. Wen, Y. Xiang, and W. Zhou, "Twitter spam detection: Survey of new approaches and comparative study," Comput. Secur., vol. 76, pp. 265284, Jul. 2018.
- [5] S. J. Soman, "A survey on behaviors exhibited by spammers in popular social media networks," in Proc. Int. Conf. Circuit, Power Comput. Technol. (ICCPCT), Mar. 2016, pp. 16.
- [6] A. Gupta, H. Lamba, and P. Kumaraguru, "1.00 per RT #BostonMarathon #prayforboston: Analyzing fake content on Twitter," in Proc. eCrime Researchers Summit (eCRS), 2013, pp. 112