ZKG INTERNATIONAL

ISSN: 2366-1313

A HYBRID APPROACH FOR DETECTING AUTOMATED SPAMMERS IN TWITTER

CH.SOWJANYA[1], T.TEJASWINI[2], V.PRANATHI3, V.GUNA SANKAR[4], K.VINODH KUMAR[5].

[1]Associate Professor, CSE,Chalapathi Institute of Technology,Guntur, India
[2]UG Student,CSE,Chalapathi Institute of Technology,Guntur, India
[3]UG Student,CSE,Chalapathi Institute of Technology,Guntur, India
[4]UG Student,CSE,Chalapathi Institute of Technology,Guntur, India
[5]UG Student,CSE,Chalapathi Institute of Technology,Guntur, India

ABSTRACT: Twitter is one of the most popular microblogging services, which generally used to share messages restricted to 280characters. However, its open nature and large user base are frequently exploited by automated spammers, content polluters, and other ill-intended users to commit various cyber crimes. Accordingly, number of approaches has been proposed by researchers to address these problems. Here in this approach we use three classifiers namely random forest, decision tree and Bayesian network .Community based features are determined to be the most effective for spam detection.

1. INTRODUCTION

Twitter, a micro blogging service, is considered a popular online social network (OSN) with a large user base and is attracting users from different walks of life and age groups. OSNs enable users to keep in touch with friends, relatives, family members, and people with similar interests, profession, and objectives. In addition, they allow users to interact with one another and form communities. A user can become a member of an OSN by registering and providing details, such as name, birthday, gender, and other contact information. Although a large number of OSNs exist on the web, Facebook and Twitter are among the most popular OSNs and are included in the list of the top 10 websites1 around the worldwide.

OSN and the Social Spam Problem Twitter, which was founded in 2006, allows its users to post their views, express their thoughts, and share news and other information in the form of tweets that are restricted to 280 characters. Twitter allows the users to follow their favourite politicians, athletes, celebrities, and news channels, and to subscribe to their content without any hindrance. Through following activity, a follower can receive status updates of subscribed account. Although Twitter and other OSNs are mainly used for various benign purposes, their open nature, huge user base, and real-time message proliferation have made them lucrative targets for cyber criminals and social bots. OSNs have been proven to be incubators for a new breed of complex and sophisticated attacks and threats, such as cyberbullying, misinformation diffusion, stalking, identity deception, radicalization, and other illicit activities, in addition to classical cyber attacks, such as spamming, phishing, and drive by download [1], [2]. Over the years, classical attacks have evolved into sophisticated attacks to evade detection mechanisms. A report2 submitted to the US Securities and Exchange Commission in August 2014 indicates that approximately 14% of Twitter accounts are actually spambots and approximately 9.3% of all tweets are spam. In social networks, spambots are also known as social bots that mimic human behaviour to gain trust in a network and then exploit it for malicious activities [3].

ISSN: 2366-1313

Such reports and findings demonstrate the extent of cyber crimes committed by spam bots and how OSNs are proving to be a heaven for these bots. Although spammers are less than benign users, they are capable of affecting network structure and trust for various illicit purposes.

2. LITERATURE SURVEY

Significant work has been done by Alex Hai Wang [1] in the year2010 which used user based as well as content based features for detection of spam profiles. A spam detection prototype system has been proposed to identify suspicious users in Twitter. A directed social graph model has been proposed to explore the "follower" and "friend" relationships. Based on Twitter's spam policy, content-based features and user-based features have been used to facilitate spam detection with Bayesian classification algorithm. Classic evaluation metrics have been used to compare the performance of various traditional classification methods like Decision Tree, Support vector Machine (SVM), Naïve Bayesian, and Neural Networks and amongst all Bayesian classifier has been judged the best in terms of performance. Over the crawled dataset of 2,000 users and test dataset of 500 users, system achieved an accuracy of 93.5% and 89% precision. Limitation of this approach is that is has been tested on very less dataset of 500 users by considering their 20 recent tweets.

In year 2010, Lee et al.[2]deployed social honeypots consisting of genuine profiles that detected suspicious users and its bot collected evidence of the spam by crawling the profile of the user sending the unwanted friend requests and hyperlinks in MySpace and Twitter. Features of profiles like their posting behaviour, content and friend information to develop a machine learning classifier have been used for identifying spammers. After analysis profiles of users who sent unsolicited friend requests to these social honey pots in MySpace and Twitter have been collected. LIBSVM classifier has been used for identification of spammers. One good point in the approach is that it has been validated on two different combinations of dataset – once with 10% spammers+90% non spammers and again with 10% non-spammers+90% spammers. Limitation of the approach is that less dataset has been used for validation.

Similarly Benevenu to et al. [3] detected spammers on the basis of tweet content and user based features. Tweet content attributes used are – number of hash tags per number of words in each tweet, number of URLs per word, number of words of each tweet, number of characters of each tweet, number of URLs in each tweet, number of hashtags in each tweet, number of numeric characters that appear in the text, number of users mentioned in each tweet, number of times the tweet has been retweeted.Fraction of tweets containing URLs, fraction of tweets that contains spam words, and average number of words that are hash tags on the tweets are the characteristics that differentiate spammers from non-spammers. Dataset of 54 million users on Twitter has been crawled with 1065 users manually labelled as spammers and non-spammers.

3. EXISTSING SYSTEM

Wang used content- and graph-based features to classify malicious and normal profiles on Twitter. In contrast to honey profiles, Wang used Twitter API to crawl the dataset.

Zhu et al. used a matrix factorization technique to find the latent features from the sparse activity matrix and adopted social regularization to learn the spam discriminating power of the classifier on the Renren network, one of the most popular OSNs in China.

Another spammer detection approach in social media was proposed by Tan et al..

ISSN: 2366-1313

This approach emphasizes the original content of genuine users that was hacked by spammers and injected with malicious links to deceive the traditional keyword- and sentence-based spammer detection techniques.

Disadvantages Of Existing System

These strategies are formal detection approaches, automated spammers can evade them by creating sufficient attack links (edges) between normal and malicious users.

Most of these approaches are based on user characterization and completely disregarding mutual interactions.

4. PROPOSED SYSTEM:

A novel study that uses community-based features with other feature categories, including metadata, content, and interaction, for detecting automated spammers A detailed analysis of the working behavior of automated spammers and benign users with respect to newly defined features.

In this study, we propose a hybrid approach for detecting social spambots in Twitter, which utilizes an amalgamation of metadata-, content-, interaction-, and community-based features.

A thorough analysis of the discriminating power of each feature category in segregating automated spammers from benign users.

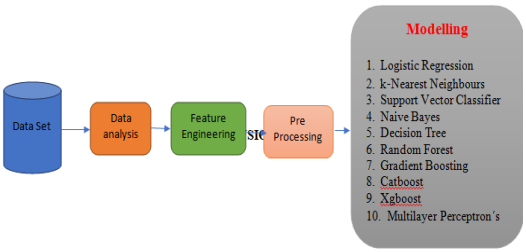
ADVANTAGES OF PROPOSED SYSTEM

The proposed approach outperforms for all the three classifiers and difference is significant for decision tree.

F Score too the proposed approach outperforms for all three classifiers.

It has shown comparatively better performance.

5. ARCHITECTURE DIAGRAM



6. IMPLEMENTATION

Data Set:

This is the step where we collect the raw data. CSV (Comma Seperated Values): A CSV file is a text file that has a specific format which allows data to be saved in a table structured format.

Data Cleaning

It is a technique which is used to transfer the raw data into a useful or efficient format such that to avoid duplication and to reduce the size of the rows.

Training the Data : The main purpose is to search the some important information in the raw data .We have used neural network technologies for training the data. Training is nothing but feature extraction.

Model Training

In any Machine Learning process, Data Preprocessing and Training is that step in which the data gets transformed, to bring it to such a state that now the machine can easily parse it. In other words, the features of the data can now be easily interpreted by the algorithm.

Ml Models

Here We are applied Various Machine learning algorithms applied. Such as

Naive Bayes

Multinomial NB

Bernoulli NB

7. CONCLUSION

In this project, Attaining perfect accuracy in spammers detection is extremely difficult, and accordingly any feature set can never be considered as complete and sound, as spammers keep on changing

ISSN: 2366-1313

their operating behaviour to evade detection mechanism. Therefore, in addition to profile-based characterization, complete logs of spammers starting from their entry in the network to their detection, need to be analyzed to model the evolutionary behavior and phases of the life-cycles of spammers. But, generally spammers are detected when they are at very advanced stage, and it is difficult to get their past logs data. Moreover, it may happen that a user is operative in the network as a benign user, and later on, it starts illicit activities due to whatsoever reasons, and considered as spammer. In this circumstance, even analyzing log data may lead to wrong characterization. Analysis of spammer's network to unearth different types of coordinated spam campaigns. Moreover, analyzing the temporal evolution of spammers' followers may reveal some interesting patterns that can be utilized for spammer's characterization at different levels of granularity.

REFERENCES

[1] .Tsikerdekis, "Identity deception prevention using common contribution network data," IEEE Transactions on Information Forensics and Security, vol. 12, no. 1, pp. 188–199, 2017.

[2] T. Anwar and M. Abulaish, "Ranking radically influential web forum users," IEEE Transactions on Information Forensics and Security, vol. 10, no. 6, pp. 1289–1298, 2015.

[3] N. R. Amit A Amleshwaram, S. Yadav, G. Gu, and C. Yang, "Cats: Characterizing automation of twitter spammers," in Proc. COMSNETS, Banglore, 2013, pp. 1–10.

[4] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "Design and analysis of social botnet," Computer Networks, vol. 57, no. 2, pp. 556– 578, 2013.

[5] D. Fletcher, "A brief history of spam," TIME, Tech. Rep., 2009.

[6] K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: Social honeypots + machine learning," in Proc. SIGIR, Geneva, 2010, pp. 435– 442.

[7] Jegadeesan,R.,Sankar Ram M.Naveen Kumar JAN 2013 "Less Cost Any Routing With Energy Cost Optimization" International Journal of Advanced Research in Computer Networking,Wireless and Mobile Communications.Volume-No.1: Page no: Issue-No.1 Impact Factor = 1.5

[8] Jegadeesan,R.,Sankar Ram, R.Janakiraman September-October 2013

[9] "A Recent Approach to Organise Structured Data in Mobile Environment" R.Jegadeesan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (6) ,Page No. 848-852 ISSN: 0975-9646 Impact Factor:2.93

[10] Jegadeesan,R., Sankar Ram October - 2013 "ENROUTING TECHNICS USING DYNAMIC WIRELESS NETWORKS" International Journal of Asia Pacific Journal of Research Ph.D Research Scholar 1, Supervisor2, VOL -3 Page No: Print-ISSN-2320-5504 impact factor 0.433

[11] Jegadeesan,R., Sankar Ram, M.S.Tharani (September-October, 2013)

[12] "Enhancing File Security by Integrating Steganography Technique in Linux Kernel" Global journal of Engineering,Design& Technology G.J. E.D.T., Vol. 2(5): Page No:9-14 ISSN: 2319 – 7293

[13] Ramesh,R., Vinoth Kumar,R., and Jegadeesan,R., January 2014

[14] "NTH THIRD PARTY AUDITING FOR DATA INTEGRITY IN CLOUD" Asia Pacific Journal of Research Vol: I Issue XIII, ISSN: 2320-5504, E-ISSN2347-4793 Vol: I Issue XIII, Page No: Impact Factor:0.433

[15] Vijayalakshmi, Balika J Chelliah and Jegadeesan,R., February-2014

[16] "SUODY-Preserving Privacy in Sharing Data with Multi-Vendor for Dynamic Groups" Global journal of

ISSN: 2366-1313

Engineering,Design& Technology. G.J. E.D.T.,Vol.3(1):43-47 (January-February, 2014) ISSN: 2319 –7293