



ENHANCE AWARENESS OF DRIFTED SPAM USING SOCIAL MEDIA

Ms. M.YESVANTA
M.E., CSE

Sri Venkateshwara College of Engineering and Technology
Thiruvallur, India.

Email: yesvantamrose@gmail.com

Mr. R. PREM KUMAR
Assistance Professor, CSE

Sri Venkateshwara College of Engineering and Technology
Thiruvallur, India

Email: premkumar.sys82@gmail.com

Abstract—

Spam has become a critical problem nowadays. In Twitter, user has follow and followers and they update the lot of tweets to dataset. The dataset is nothing but set of data files that is more than one file stored in one file. In our Tweets or our follow tweets sometimes there will be a Spam tweets. He or She can update their tweets in their wall. They can share their tweets by text or links. Links also they update. Links can be both spam tweets and non-spam tweets. Spam is nothing but unnecessary tweets or links based. Tweets are nothing but sharing textual information or message. Using that spam, spam providers can easily get our identities like username and password etc. When one he or she updates the spam links, the other follower can view that spam tweets. If they click spam detection will detect the spam and alert the user to not open the link. This helps the follows to safe their details from spam tweets. Twitter spam drift problem through analyzing the statistical properties of Twitter spam in our collected dataset and then its impact on detection performance of using random forest classifier algorithm. In order to tackle this problem, we propose Random Forest Classifier algorithm to analyze and detect the large number of spam and non-spam tweets by users we firstly carry out a deep analysis on the statistical features of one million spam tweets and one million non-spam tweets

I. INTRODUCTION

A spam detection

A spam filter is a program that is used to detect

Unsolicited and unwanted email and prevent those messages from getting to a user's inbox. Like other types of filtering programs, a spam filter looks for certain criteria on which it bases judgments. [3] discussed about creating Obstacles to Screened networks. In today's technological world, millions of individuals are subject to privacy threats. Companies are hired not only to watch what you visit online, but to infiltrate the information and send advertising based on your browsing history. People set up accounts for facebook, enter bank and credit card information to various websites. Those concerned about Internet privacy often cite a number of privacy risks events that can compromise privacy which may be encountered through Internet use. These methods of compromise can range from the gathering of statistics on users, to more malicious acts such as the spreading of spyware and various forms of bugs (software errors) exploitation.

Spamming

Spam is email spam, the term is applied to similar abuses in other media: instant messaging spam, Usenet newsgroup spam, Web search engine spam, spam

Random Forest Classifier Algorithm

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression).



II. RELATEDWORKS

Modules means emphasizes separating the functionality of a program into an independent interchangeable, such that each contains everything necessary to execute only one aspect of the desired functionality. This is the stage of the project when the theoretical design is turned out in to a working system.

A. FOLLOWS

Here, the user can follow their friends or relatives by searching their profiles in website. By this following user can be able to view the all details of other users those who followed by him/her. In which only authentication person are followed and share their tweets or information in the social media application.

B. FOLLOWERS

The user was followed by another user. User can be able to view the all details of that user who followed this user. In which follows and follower can share their own details or information in interference of application. So, the request has send from follower and then the follower response is received from follows by authorized admin for follows and follower interaction. The tweets or sharing messages or information are properly and maintained by authorized controller.

C. PROFILEUPDATES

Here, user can update their profiles which are DOB, places, images etc. This helps the other find out the user details. We can update the details daily when we want to update the details. In which the user profiles are maintained by application controller. The application controller maintains the user tweets, profiles, etc.

D. TWEETS

User can update their tweets in their wall. They can share their tweets by text or links. Links also they update. Links can be both spam tweets and non-spam tweets. Spam Tweets is nothing but URL based. Spam tweets URLs is duplicate website of original websites.

E. SPAM DETECTION

The spam detection is fully responsible for admin. The authorized controller roles are maintaining

user profiles, user tweets and alert spammed link from user. When one user updates the spam links, the other user can view that spam tweets. If they click spam detection will detect the spam and alert the user to not open the link. This helps the user to safe their details from spam tweets

III. EXISTINGSYSTEM

Twitter has become one of the most popular social networks in the last decade. It is rated as the most popular social network among teenagers according to a recent report. However; the exponential growth of Twitter also contributes to the increase of spamming activities. Twitter spam, which is referred to as unsolicited tweets containing malicious link that directs victims to external sites containing malware downloads, phishing, drug sales, or scams, etc. not only interferes user experiences, but also damages the whole Internet. In which lots of users suffer from the problem of "Spam Drift" due to the change of statistical features of spam tweets as time goes on.

When "spam drifts", the old classification model such as phishing, scams, and software downloading is not updated with "changed" spam samples, as a result, the classification results will gradually become inaccurate.

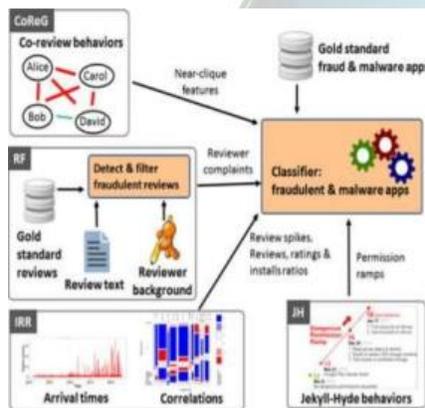
IV. PROPOSEDSOLUTION

Existing machine learning based spam detection methods suffer from the problem of "Spam Drift" due to the change of statistical features of spam tweets as time goes on. When "spam drifts", the old classification model is not updated with "changed" spam samples, as a result, the classification results will Gradually become inaccurate. To solve this problem, obtaining the "changed" samples to update the classification model is very important. By observing that there are such Samples in the unlabeled incoming tweets which are very easy to collect, we propose a scheme called "Lfun" to address "Spam Drift" problem. This section presents our Lfun scheme to deal with the drift problem in Twitter spam detection. There are two main components in this framework: LDT is to learn from detected spam tweets and LHL is to learn from human labelling. In "Drifted Spam Detection" scenario, we have already got a small amount of labelled spam and non-spam tweets.

V. SYSTEMDESIGN

System design involves identification of classes their relationship as well as their collaboration. In object-oriented design, classes are divided into entity classes and control classes. The Computer Aided Software Engineering (CASE) tools that are available commercially do not provide any assistance in this transition. CASE tools take advantage of Meta modeling that are helpful only after the construction of this diagram. In the FUSION method, some object-oriented approach like Object Modeling Technique (OMT), classes, and Responsibilities. Any software project is worked out by both the analyst and the designer. The analyst creates the user case diagram. The designer creates the class diagram.

System Architecture



A system architecture or systems architecture is the conceptual design that defines the structure and behavior of a system. An architecture description is a formal description of a system, organized in a way that supports reasoning about the structural properties of the system. It defines the system components or building blocks...and provides a plan from which products can be procured, and systems developed, that will work together to implement the overall system.

System architecture. The CoReG module identifies suspicious, time related co-review behaviors. The RF module uses linguistic tools to detect suspicious behaviors reported by genuine

reviews. The IRR module uses behavioral information to detect suspicious apps. The JH module identifies permission ramps to pinpoint possible Jekyll-Hyde app transitions.

VI. CONCLUSION

In future, we can easily focus and avoid on spam tweets with URLs. To solve this problem, obtaining the “changed” samples to update the classification model is very important.

By observing that there are such samples in the unlabeled incoming tweets which are very easy to collect, we propose a scheme called “Random Forest Classifier” to address “Spam Drift” problem.

REFERENCES

- [1] Google Play. [Online]. Available: <https://play.google.com/>
- [2] E. Siegel, “Fake reviews in Google Play and Apple AppStore,” Appnitive, Seattle, WA, USA, 2014.
- [3] Christo Ananth, P.Muppudathi, S.Muthuselvi, P.Mathumitha, M.Mohaideen Fathima, M.Muthulakshmi, “Creating Obstacles to Screened networks”, International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST), Volume 1, Issue 4, July 2015, pp:10-14
- [4] S. Mlot. (2014, Apr. 8). “Top Android App aScam, Pulled From Google Play,” PCMag. Available: <http://www.pcmag.com/article2/0,2817,2456165,00.asp>
- [5] D. Roberts. (2015, Jul. 8). “How to spot fake apps on the Google Play store,” Fortune. Available: <http://fortune.com/2015/07/08/Google-play-fake-app/>
- [6] A. Greenberg (2012, May 23). “Researchers say they snuck Malware app past Google’s ‘Bouncer’ Android market scanner,” Forbes Security, [Online]. Available: <http://www.forbes.com/Sites/andygreenberg/2012/05/23/researchers-say-they-snuckmalware-app-past-googles-bouncer-android-market-scanner/#52c8818d1041>
- [7] Freelancer. [Online]. Available: <http://www.freelancer.com>
- [8] Fiverr. [Online]. Available: <https://www.fiverr.com/>
- [9] BestAppPromotion. [Online]. Available: www.bestreviewapp.com