Spam Detection Framework Based on Text Reviews in Online Social Networks

INDUKURI MOUNICA SRI [#1] & Y.SRINIVASA RAJU [#2] & D.D.D.SURIBABU [#3],

[#1] MCA Student, Master of Computer Applications, D.N.R. College, P.G.Courses & Research Center, Bhimavaram, AP, India.

[#2] Assistant Professor, Master of Computer Applications, D.N.R. College, P.G.Courses & Research Center, Bhimavaram, AP, India.

[#3] Head & Associate Professor, Dept of CSE, D.N.R. College of Engineering, Bhimavaram, AP, India.

Received: January 17, 2019 Accepted: February 21, 2019

ABSTRACT: *In recent days, almost all the people mainly rely mostly on social media for communicating with others. They almost depend on OSN to take decision about any product or topic based on the reviews. As we all know that lot of people will give different reviews based on their individual interest and this will leave an opportunity for the spam users to leave a spam review for those products. Identifying these spammers and the spam content is a hot topic of research and although a considerable number of studies have been done recently toward this end, but so far the methodologies put forth still barely detect spam reviews, and none of them show the importance of each extracted feature type. In this application we try to identify the spam detection based on the reviews given by the end users based on some keywords. By conducting various experiments on our proposed model we finally came to an conclusion that our proposed approach is best in identifying the spam users from a set of reviews that is posted based on some spam keywords.The simulation results clearly tells that our application is best in separating the spam accounts and normal accounts separately into two lists which is not their in current OSN networks.*

Key Words*: Spam Keywords, Reviews, Online Social Network (OSN), Spammers, Framework.*

I. Introduction

In recent days internet has gained its momentum in attracting a lot of users towards it with the help of several social networking sites[1]. One among the several OSN is facebook in which the facebook is one of the major source for communication in a very fast manner. Recently in the osn network we can see a lot of spammers try to give wrong feedbacks and reviews for the products or posts which is shared among various users[2]. This is not at all identified or blocked by any administrator as we can't able to identify who is giving the abused messages in form of comments or replies to the OSN network. There is no pre-defined method to identify and block such a spam users not to post such messages on genuine user walls.This motivated us to design this current application in which network based spam detection can be identified based on some spam keywords which is posted on the user OSN walls[3].

Figure 1. Demonstrate the Sample Flow of Online Social Networks

From the above figure 1, we can clearly identify that online social network is a collection of various resources like web, media, link, connect and so on. This is mainly used in order to connect one person with other around the world and share their updates through internet. During the communication they may share their personal and emotional feelings and they may expect replies from others. During this comments and replies time there are several problems that takes place while we try to post the reviews and view the reviews. The main problem in current days review system is there is no concept like identifying the spam users reviews and normal reviews separately[4].

II. Related Work

In this section we will mainly discuss about the background work that was carried out in order to prove the performance of our proposed model to identify the spam users who try to create negative impact on the user posts.

Motivation

The main motivation for writing up this paper is as follows:

Here we try to examine the evidence of the existing manipulation of OSN topics. In particular, employing an influence model, we analyze the dynamics of an endogenous hash tag and identify the manipulation from its endogenous diffusion. After further investigating the manipulation in the dynamics, we disclose the existence of a suspect spamming infrastructure.

We study the OSN Post at topic level, considering topics' popularity, coverage, transmission, potential coverage, and reputation. The corresponding dynamics for each factor above are extracted, and then Support Vector Machine (SVM) classifier is used to check how accurately a factor could predict the accuracy of spam[6]. We find that, except for transmission, each studied factor is associated with osn post. We further illustrate the interaction pattern between malicious accounts and authenticated accounts, with respect to posted message in OSN.

Here we try to present a list of normal messages for the User OSN wall and also try to present a set of malicious messages for the messages which is posted based on spam content. Here in order to show the difference of both the types of messages, we kept normal messages as one list and malicious or spam messages as another list. Here the malicious accounts are one which doesn't have any account [7] and they try to login and post different comments which is not at all related to that current topic and these type of tweets or messages should be identified immediately and blocked at the system level[8].

III. Proposed Spam Detection Framework Based on Text Reviews in Online Social Networks

In this section we will mainly discuss about the proposed Spam Detection Framework Based on Text Reviews in Online Social Networks. Now let us discuss about this proposed model in detail as follows:

Preliminary Knowledge

In this proposed system we try to use NetSpam a novel feature to find features importance even without ground truth, and only by relying on metapath definition and based on values calculated for each review[9]. NetSpam improves the accuracy compared to the state of- the art in terms of time complexity, which highly depends to the number of features used to identify a spam review; hence, using features with more weights will resulted in detecting fake reviews easier with less time complexity. This proposed netspam can easily identify the spam reviews and normal reviews easily based on the text reviews[10].

Advantages of our Proposed System

The following are the advantages of our proposed approach. They are as follows:

1. To identify spam and spammers as well as different type of analysis on this topic.
2. Written reviews also help service providers to enhance the quality of their products and services.
3. By using this proposed technique we can able to identify the spam reviews based on the keywords that are posted in the reviews
4. Here the system will automatically identify the reviews and automatically identifies the spam users and try to tag the user with spam label once based on their reviews.

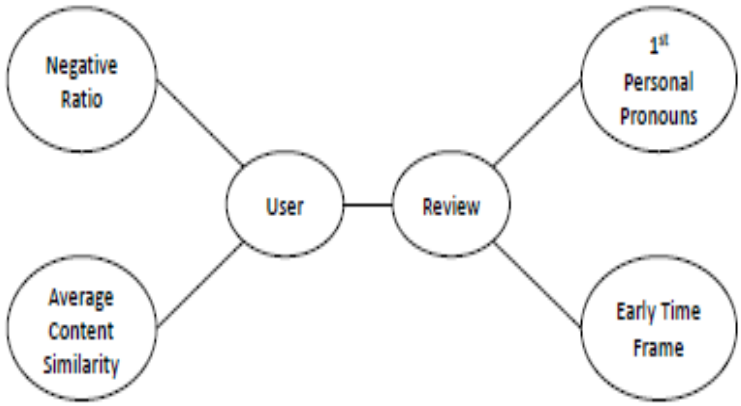

Figure 2. Represents Proposed Architecture to find out the Average Ratio and Content Similarity of Normal Reviews with Spam Reviews

From the above figure 2,we can clearly identify the difference in the average content similarity as well as negative ration with content similarity that is available based on the user review[12].Here we try to identify the review based on the keywords that are matched with some spam keywords and if any message is matched with the spam keyword that is treated as spam message and if any message is not matched with any spam keywords and that is coming from a valid accounts then they will be treated as normal account and normal message.In this way we can able to separate the messages into two separate categories like spam and not spam[11].

IV. Implementation Phase

Implementation is a stage where the theoretical design is converted into programmatically manner and in this stage we try to divide the application into number of modules. In our proposed application we try to divide into two modules.

1) Admin Module

In this module, the Admin has to login by using valid user name and password. After login successful he can do some operations such as adding Categories, Adding Products for that Categories, Viewing and authorizing users, View Spam accounts details,Viewing friend request & response, All recommended posts, All posts with all Reviews, All Positive and Negative Reviews , Removing Products, Viewing All Purchased Products, viewing Positive and Negative Reviews Chart on products.

2) User Module

In this module, there are n numbers of users are present. User should register before doing 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 will do some operations like viewing their profile Account details like Spam or Normal , search users and send friend request, viewing friend requests, searching posts and recommend to friends and viewing all product recommendations sent to him by his friends, commenting on posts, purchasing products and viewing their product search history.

V.Conclusion

In this paper, we for the first time design and implement the spam detection based on the reviews given by the end users based on some keywords. By conducting various experiments on our proposed model we finally came to an conclusion that our proposed approach is best in identifying the spam users from a set of reviews that is posted based on some spam keywords. The simulation results clearly tells that our application is best in separating the spam accounts and normal accounts separately into two lists which is not there in current OSN networks.

VI. References

1. M. Ott, Y. Choi, C. Cardie, and J. T. Hancock. Finding deceptive opinion spam by any stretch of the imagination.In ACL, 2011.
2. Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. In SIAM International Conference on Data Mining, 2014.

3. A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. Trueview: Harnessing the power of multiple review sites. In ACM WWW, 2015.
4. B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Towards detecting anomalous user behavior in online social networks. In USENIX, 2014.
5. N. Jindal and B. Liu. Opinion spam and analysis. In WSDM, 2008.
6. F. Li, M. Huang, Y. Yang, and X. Zhu. Learning to identify review spam. Proceedings of the 22nd International Joint Conference on Artificial Intelligence; IJCAI, 2011.
7. G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Exploiting burstiness in reviews for review spammer detection. In ICWSM, 2013.
8. H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.
9. L. Akoglu, R. Chandy, and C. Faloutsos. Opinion fraud detection in online reviews bynetwork effects. In ICWSM, 2013.
10. R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review networksand metadata. In ACM KDD, 2015.
11. S. Feng, R. Banerjee and Y. Choi. Syntactic stylometry for deception detection. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers; ACL, 2012.
12. N. Jindal, B. Liu, and E.-P. Lim. Finding unusual review patterns using unexpected rules. In ACM CIKM, 2012.