NETSPAM: A NETWORK-BASED SPAM DETECTION FRAMEWORK FOR REVIEWS IN ONLINE SOCIAL MEDIA

[1]V.Muniraj Naidu, , [2]C.Dinesh, [3]Dr.B.Sunil Kumar

[1]Associate Professor, [2]Assistant Professor, [3]professor
Dept. of CSE,
Narayana Engineering College, Nellore, A.P, India

Abstract- Nowadays, a big part of people rely on available content in social media in their decisions (e.g. reviews and feedback on a topic or product). The possibility that anybody can leave a review provide a golden opportunity for spammers to write spam reviews about products and services for different interests. 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 study, we propose a novel framework, named NetSpam, which utilizes spam features for modeling review datasets as heterogeneous information networks to map spam detection procedure into a classification problem in such networks. Using the importance of spam features help us to obtain better results in terms of different metrics experimented on real-world review datasets from Yelp and Amazon websites. The results show that NetSpam outperforms the existing methods and among four categories of features; including review-behavioral, user-behavioral, review linguistic, user-linguistic, the first type of features performs better than the other categories..

Keywords- NetSpam, Heterogeneous Information Network (HIN)

1.INTRODUCTION

Online Social Media portals play an influential role in information propagation which is considered as an important source for producers in their advertising campaigns as well as for customers in selecting products and services. In the past years, people rely a lot on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of products and services. In addition, written reviews also help service providers to enhance the quality of their products and services. These reviews thus have become an important factor in success of a business while positive reviews can bring benefits for a company, negative reviews canpotentially impact credibility and cause economic losses. The fact that anyone with any identity can leave comments as review provides a tempting opportunity for spammers to write fake reviews designed to mislead users' opinion. These misleading reviews are then multiplied by the sharing function of social media and propagation over the web. The reviews written to change users' perception of how good a product or a service are considered as spam, and are often written in exchange for money.As shown, 20% of the reviews in the Yelp website are actually spam reviews. On the other hand, a considerable amount of literature has been published on the techniques used to identify spam and spammers as well as different type of analysis. These techniques can be classified into different categories; some using linguistic patterns in text which are mostly based on bigram, and unigram, others are based on behavioural patterns that rely on features extracted from patterns in users' behaviour which are mostly metadata based and even some techniques using graphs and graph-based algorithms and classifiers.

Despite this great deal of efforts, many aspects have been missed or remained unsolved. One of them is a classifier that can calculate feature weights that show each feature's level of importance in determining spam reviews. The general concept of our proposed framework is to model a given review dataset as a Heterogeneous Information Network (HIN) and to map the problem of spam detection into a HIN classification problem. In particular, we model review dataset as a HIN in which reviews are connected through different node types (such as features and users). A weighting algorithm is then employed to calculate each feature's importance (or weight).These weights are utilized to calculate the final labels for reviews using both unsupervised and supervised approaches.To evaluate the proposed solution, we used two sample review datasets from Yelp and Amazon websites. Based on our observations, defining two views for features (review-user and behavioural-linguistic), the classified features as review behavioural

have more weights and yield better performance on spotting spam reviews in both semi- supervised and unsupervised approaches. In addition, we demonstrate that using different supervisions such as 1%, 2.5% and 5% or using an unsupervised approach, make no noticeable variation on the performance of our approach. We observed that feature weights can be added or removed for labelling and hence time complexity can be scaled for a specific level of accuracy. As the result of this weighting step, we can use fewer features with more weights to obtain better accuracy with less time complexity. In addition, categorizing features in fourmajor categories (review-behavioural, user-behavioural, review linguistic, user-linguistic), helps us to understand how much each category of features is contributed to spam detection. In summary, our main contributions are as follows:We propose Net Spam framework that is a novel network based approach which models review networks as heterogeneous information networks. The classification step uses IEEE Transactions on Information Forensics and Security, Volume: 12, Issue: 7, Issue Date: July.2017 2 different Meta path types which are innovative in the spam detection domain.A new weighting method for spam features is proposed to determine the relative importance of each feature and shows how effective each of features are in identifying spams from normal reviews. Previous works also aimed to address the importance of features mainly in term of obtained accuracy, but not as a build-in function in their framework (i.e., their approach is dependent to ground truth for determining each feature importance). As we explain in our unsupervised approach, Net Spam is able to find features importance even without ground truth, and only by relying on meta path definition and based on values calculated for each review.Net Spam 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.**II.**

II.LITERATURE REVIEW

The purpose of Information Security is to protect the valuable resources of an organization such as hardware, software and skilled people. Through the selection and application of appropriate safeguard, security helps the organization to meet its business objectives or mission by protecting its physical and financial resources, reputation, legal position, employees and other tangible and intangible assets..This Researcher has collected comprehensive information from various books, manuals, magazines, journals, articles and research websites. Information gathered through various seminars and conferences attended by her also helped lot for the researcher.

ELEMENTS OF INFORMATION SECURITY: Thomas R. Peltier who is an Information security professional since 1977, has provided guidelines for effective Information Security Management. Referring to his guidelines, information security should be based on following eight major elements

a.   Information protection should support objectives of business or mission of enterprise. Many times information security personnel lose track of their goals and responsibilities. The post of Information Security Officer has to be created to support enterprise security.
b.   Information protection is an integral element for necessary care. Senior management is endowed with two basic responsibilities such as duty of loyalty and duty of care. A duty of loyalty implies that the decisions shall be made in the best interest of the enterprise and duty of care implies that senior management shall protect the assets of the enterprise and make informed business decisions.
c.   Information protection must be cost-effective. Implementation of controls must be proposed and it is necessary to confirm that a significant risk exists. Implementing a timely risk analysis process can accomplish   t

d. Information security responsibilities and accountabilities should be made explicit. For any program to be more effective, it is necessary to publish information security policy statement and information security group mission statement. The information security policy should identify roles and responsibilities of all employees. To make the policy more effective the language of the policy must be incorporated into purchase agreements for all contract personnel and consultants

e. Systems owners have information protection responsibilities within the own organization. Access to information often extends beyond the organization also. This is the responsibility of information owner. The main responsibility is to monitor the usage to ensure that it complies with the user profile and authorization of the users. If system has external users then the owners have the responsibility to share appropriate level of knowledge about the existence and general extent of control measures so that other users can be confident that the system is adequately secured.

f. Information protection requires a comprehensive and integrated approach. For its effectiveness, it is necessary that the information security and its protection shall be a part of the system development lifecycle. During initial or analysis phase, information security should include risk analysis, a business impact analysis andinformation classification document. Additionally, because information is resident in all departments throughout the enterprise, each business unit should establish an individual responsible for implementing the information protection program to meet the specific business goals of the department.

g. Information security should be periodically reassessed with respect to time, need and objectives. A good information protection program examines itself on a regular basis and makes changes wherever and whenever necessary. This should be a dynamic process hence must be reassessed every eighteen months or even earlier on extraordinary occasions.

h. Information protection is constrained by culture of the organization. The information System Security Officer must understand and advice basic information security program that must beimplementedthroughouttheenterprise.

III. SECURITY PRINCIPLES:

Following are some of the guidelines issued by Economic Cooperation and

Development intended for development of laws and policies:

a. **Accountability**: Everybody involved with the security of information must have specified accountability towards actions.

b. **Awareness:** Everyone from the organization must be able to access the knowledge pertaining to security measures, practices and procedures and all efforts shall be made in building confidence in information systems.

c. **Ethics:** The method in which informationsystems and their associated security mechanisms are operated must be able to respect the privacy, rights and legitimate interests of users.

d. **Multidisciplinary principle**: All the aspects and opinion must be considered in the development of policies, procedures and techniques. These must include legal, technical, administrative, organizational, operational, commercial and educational aspects.

e. **Proportionality**: Security measures must be based on the value of information and the level of risk involve

f. **Integration:** Security measures must be integrated to work together and establish defensive depth in the security system.

g. **Timeliness**: Everyone should act together in coordinated and timely fashion when a security breach occurs.

h. **Reassessment:** Security mechanisms and needs must be reassessed periodically to ensure that organizations needs are being met.

IV.SYSTEM ANALYSIS

The results show that Net Spam outperforms the existing methods and among four categories offeatures; including review-behavioural, use behavioural, review linguistic, user linguistic. The first type of features performs better than the other categories.Despite this great deal of efforts, many aspects have been missed or remained unsolved. One of them is a classifier that can calculate feature weights that show each feature's level of importance in determining spam reviews. The general concept of our proposed framework is to model a given review dataset as a Heterogeneous Information Network (HIN) and to map the problem of spam detection into a HIN classification problem. In particular, we model review dataset as a HIN in which reviews are connected through different node types. The general concept of our proposed framework is to model a given review dataset as a Heterogeneous Information Network and to map the problem of spam detection into a HIN classification problem. In particular, we model review dataset as in which reviews are connected through different node types. A weighting algorithm is then employed to calculate each feature's importance. These weights are utilized to calculate the final labels for reviews using both unsupervised and supervised approaches.

MODULES

Admin
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.

Search Users The user can search the users based on names and the server will give response to the user like User name, user image, E mail id, phone number and date of birth. If you wish to send friend request to particular user then click on "request" button, then request will be send to that particular user

.**Searching Products and Recommend to Friends**In this, the user searches for products based on the products description. The user can recommend searched products to his friends, comment on post and he can add the products to cart to buy those added products later by using their created account.

View Friend Requests In this module, the user can view the friend requests which are sent by other users. Which includes request sent user details with their tags such as user name, user image, date of birth, E mail ID, phone number and Address and user can accept the request by clicking on the "waiting" link.

View Product Recommends In this module, the user can view all the products which are recommended by his friends. This includes recommended user name and his image, recommended products details.

View Product Search History In this module, the user can view all the searched products names and categories, the keywords which he used to search the products. This includes details such as, searched product, used keyword and date of search.

View Bank Account Details In this module, the user can create his bank account by providing details such as, account number, branch, address, email id. Later he can add money to his account and can view his account details.

V.EXPERIMENTAL RESULTS
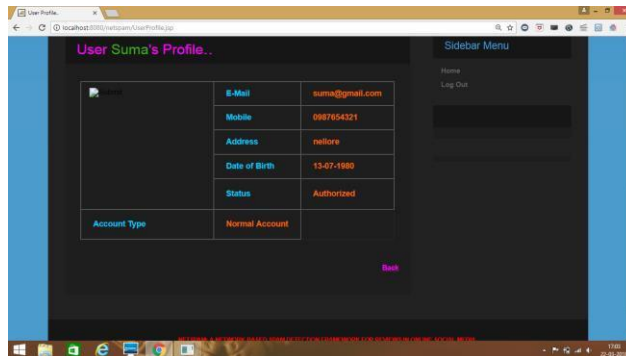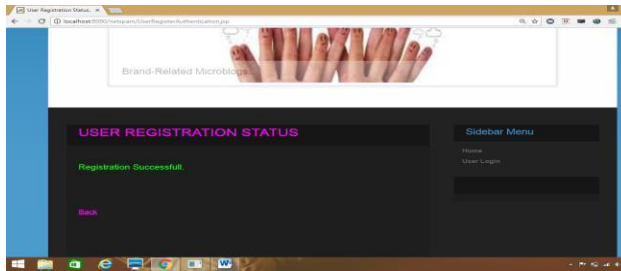
Security,Volume:12,Issue:7,Issue Date:July.201710 a new graph-based method to label reviews relying on a rank-based labeling approach. The performance of the proposed framework is evaluated by using two real-world labeled datasets of Yelp and Amazon websites. Our observations show that calculated weights by using this meta path concept can be very effective in identifying spam reviews and leads to a better performance. In addition, we found that even without a train set, Net Spam can calculate the importance of each feature and it yields better performance in the features' addition process, and performs better than previous works, with only a small number of features. Moreover, after defining four main categories for features our observations show that the reviews behavioural category performs better than other categories, in terms of AP, AUC as well as in the calculated weights. The results also confirm that using different supervisions, similar to the semi-supervised method, have no noticeable effect on determining most of the weighted features, just as in different datasets.

REFERENCES

[1] J. Donfro, A whopping 20% of yelp reviews are fake. http://www.businessinsider.com/20-percent-of-yelp-reviews-fake-2013-9.Accessed: 2015-07-30.

[2] M. Ott, C. Cardie, and J. T. Hancock. Estimating the prevalence of deception in online review communities. In ACM WWW, 2012.