Spam Review Detection Framework for Online Social Media

Anand S Hiremath[1], Heena M Sangtrash[2]

[1]Assistant Professor, Department of Computer Science and Engineering, BLDEA's V.P.Dr.P.G.Halakatti College of Engineering & Technology, Vijayapur, Karnataka, India

[2]M.Tech Student, Department of Computer Science and Engineering, BLDEA's V.P.Dr.P.G.Halakatti College of Engineering & Technology Vijayapur, Karnataka, India

ABSTRACT

The online social media helps people to buy the different products through an application, where the user can select the different products with various price. Also the user can write the review on the product and provide ratings for the product as well. As anybody can write the reviews on a product, these reviews of a product plays an essential role in the online social media, as the positive reviews may encourage the users for purchasing of the product but the negative review may have the adverse affect as it may decrease the selling of product. Many of the users buy the products on a site based on the reviews of the product believing it to be true, and many of the spammers write the fake reviews which may have bad impact of the product on the user and hence the organizations loss for not selling of the product. So it is necessary to find out whether the reviews on a product is spam or not. The spam detection framework finds the spam reviews that are written by spammers based on feature types.

Keywords: Social Media, Spam Review, Heterogeneous Information Networks.

1. INTRODUCTION

For the choice of the products and the services and also to the producer of advertising drives, data propagation is an important source by the online social media gateway. The reviews written by the user on the products affects the purchase of the products by other users. The positive review on the product encourages user to buy the product but the negative review does the vice versa where the users are discouraged and hence doesn't buy the product. These reviews will also lead the improvement of the product quality and services quality as well. The business success is dependent on these reviews which are considered as the major factors where the company gets benefits based on positive reviews and the negative reviews can cause the economically losses for the company. The spammers can write the fake reviews as they get an opportunity as anybody with an identity can write the comments as a review and which may mislead the user's opinion. The social media's sharing function then multiplies these misleading reviews. The change of user's point of view on whether the product or service is good or not based on written reviews are believed to be spam, and these are mainly written for the purpose of money. In [1], it shows that in the Yelp website the spam reviews are actually of about 20% .

Despite of these efforts, many of the aspects are missed and are unsolved. In the determination of spam reviews classifier is used which calculates feature weights which shows each features importance. The given review dataset is modeled into Heterogeneous Information Network(HIN) this is the common concept of the proposed framework and the problem of detecting the spam is plotted into the problem of HIN classification. The HIN is modeled from the review dataset where the different node types like features and users are used to connect the reviews. To determine each features importance or weight weighting algorithm is used. The supervised and unsupervised methods are used to find the review labels using these calculated weights.

For the features which are review-user and behavioral-linguistic two views are defined according to the observations, the more weight is in review behavioral feature and in both semi-supervised and unsupervised methods the spam review detection has

2. LITERATURE SURVEY

Nitin Jindal et al, [1] There are generally three types of spam reviews i.e. Brand review only which is done manually, Non-review which is also done manually and lastly Untruthful review which has three types of duplicates as follows: i) Duplicates on the same product by different user ids. ii) Duplicates on different products by the same user id. iii) Duplicates on different products by different user ids.

 DOI : 10.18231/2454-9150.2018.0241

E Lim et al, [2] for the determination of each reviewer's spam degree a method is proposed which is applied on the dataset of Amazon reviews. A subset of doubtful reviewers is choosed from the extremely doubtful reviewers for further study. The user assess the spammers using the software which is the web based, then the distrustful reviewers subset is selected for additional survey and the user evaluates these in the experiment.

A J Minnich et al, [3] proposed methodology so as to combine, compare and evaluation of reviews that are from numerous hosted sites. The emphasis is on the reviews of hotel also use 15 million reviews and more on the three prominent travel sites where more than the 3.5 million users span on it. The system has three impels firstly to effectively identify the cross-site difference the novel features are developed, secondly the huge study using the real data for the cross-site variations and then with accuracy of 93% developing a hotel identify matching, and lastly for the proof that the end user is informed with the better cross-site analysis a TrueView is introduced.

B Viswanath et al, [4] to represent the normal users behavior accurately Principal Component Analysis(PCA) technique is presented. The reaction of users that are normal is experimentally evaluated and it is in the subspace of low dimension complaint to the technique which is PCA. By using the data from Facebook the approach effectiveness is demonstrated. The strategies of various attackers is detected successfully and then the Facebook detects the labels with having false or positive rates and finally, in the ads of Facebook to identify the clickspam the proposed approach is used.

H li et al, [5] multi typed heterogeneous collective classification which is a classification is proposed by the superiority of dependencies which are complicated between the users, IP addresses and the reviews and unlabeled and the positive learning is obtained from these. In China and Shanghai the 500 restaurants are choosed and the real reviews from these are used for the experiment. As the features which are independent of language is used in the models, easily they can be unspecialized into other languages.

3. OBJECTIVE

1. The 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.
2. In particular, model a review dataset as a HIN in which reviews are connected through different node types (such as features and users)[6].
3. A weighting concept 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.

4. METHODOLOGY / PLANNED WORK

Calculations for detecting spam reviews for each of the modules are:

1)User-behavioral(UB) based

i) Burstiness: There are mainly two reasons for which the spammers in short period of a time write spam reviews: Firstly, to have a influence on readers and the other users, and secondly as the users who are temporal wants to write reviews as possible as can in a less period of interval.

3. **ii) Negative Ratio:**Collect user rating's of particular item. Average of rating score for more than two times is equal **(=)** or less than **(<)** 2 which is spam , more than 2 is not spam for: user-john gave for "prod**1**" ratings are

2) User Linguistic Based

i) Average Content Similarity:

First we need to check for a review if any other review if any other reviews are there for same product with same user then we need to check otherwise it is not spam.

3) Review Behavioral Based

i) Early Time Frame:

To let the other users to visit the spam reviews first, the spammers try to write their reviews as early as they can so that their reviews are in the top reviews..

ii) Rate Deviation

 DOI : 10.18231/2454-9150.2018.0241

Spammers want to make promotion of the business that they are in a contract with, because of which they rate those business with a high scores.

4) Review Linguistic Based

Studies shows that the personal pronouns are often used by the spammers. And also the spammers put '!' as possible as in the sentences for having more impact on the user also so that users can highlight their reviews among the other ones.

5. CONCLUSION

Spam review detection framework detects the fake reviews that are written by the spammers. The admin adds the items for purchasing. The user purchases the items and provide ratings and gives review on the product. These reviews affects the other users for buying the products, as the positive review can lead to increase selling of the product where as the negative review may lead the un-purchasing of product by others. These reviews which are negative are often written by the spammers. Hence in this spam review detection, the spam reviews are identified using the algorithm which calculates weight of each spam feature and the review labels are computed using the computed weights by the supervised and the unsupervised method.

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] G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. "Exploiting the burstiness in reviews for review spammer detection," in ICWSM, 2013.

[3] A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. "Trueview: Harnessing power of the 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 the anomalous user behavior in the online social networks," in USENIX, 2014.

[5] H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. "Spotting the fake reviews via collective PU learning," in ICDM, 2014.

[6] S Shehnepoor, M Salehi, R Farahbakhsh, N Crespi. " NetSpam: a Network-based Spam Detection Framework for Reviews in Online Social Media" in IEEE Conference 2017.

 DOI : 10.18231/2454-9150.2018.0241