PLEXUS SPAM UNMASKING SCHEMA SCRUTINY IN SOCIAL WEB

Satish KumarA[1], Prof. E. Srikanth Reddy[2]

1. **M.Tech Scholar, Department of CSE, Vaageswari College of Engineering, Karimnagar, Telangana, India–satishumar.avunoori@gmail.com**
2. **Associate Professor, Department of CSE , Vaageswari College of Engineering, Karimnagar, Telangana, India-srikanth574@gmail.com**

ABSTRACT

Nowadays, a big part of people rely on available content in social media in their decision for example, 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 Plexus Spam, 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 Plexus Spam 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.

I. 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 can potentially 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 [11], and are often written in exchange for money.

As shown in [1], 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 on this topic [30], [31]. These techniques can be classified into different categories; some using linguistic patterns in text [2], [3], [4], which are mostly based on bigram, and unigram, others are based on behavioral patterns that rely on features extracted from patterns in users' behavior which are mostly metadata based [34], [6], [7], [8], [9], and even some techniques using graphs and graph-based algorithms and classifiers [10], [11], [12].

II. EXISTING SYSTEM

The results show that Plexus Spam outperforms the existing methods and among four categories of features; including review-behavioral, use behavioral, 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.

Disadvantages of Existing System:

- This utilizes spam features for modeling review datasets as heterogeneous information networks to map spam detection procedure into a classification problem in such networks.
- Time Complexity.

III. PROPOSED SYSTEM

Plexus Spam is able to find featuresimportance even without ground truth, and only by relying on meta path definition and based on values calculated for each review. Plexus Spam improves the accuracy compared to the state of-the art in terms of time complexity, which highly depends tothe number of features used to identify a spam review; hence using features with more weights will resulted in detectingfake reviews easier with less time complexity.

A new Content Based Algorithm for spam features is proposedto determine the relative importance of each featureand shows how effective each of features are in identifyingspams from normal reviews.

Advantages of Proposed System:

- To identify spam and spammers as well as different type of analysis on this topic. Written reviews also help service providers to enhance the quality of their products and services.

IV.IMPLEMENATATION

- **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.

    Adding Categories

    In this module, the admin adds the category details such as category name. These details will be stored into the database.

    Adding Products

    In this module, the admin adds Product posts for categories which include details such as, product image, product name, cost, description and uses of that product. These details will be stored into the

database. These details will be further searched and accessed by the users in order to recommend to their friends and to buy products.

Authorize Users

In user's module, the admin can view the list of users who all registered. In this, the admin can view the users' details such as, user name, email, address, phone number and authorize the users.

Request & Response

In this module, the admin can view all the friend requests and responses.Here all the requests and responses will be displayed with their tags such as Id, requested user image, requested user name, user name request to, status and time & date. If the user accepts the request then the status will be changed to accepted or else the status will remains as waiting.

All Recommended Posts

In this module, the admin can view all the recommended products. If any recommendations happened for particular products, those details will be shown along with products. Details include product name, recommended user name, user recommended to name and the date.

View Positive /Negative Comments

In this, the admin can view all posts with their Positive and Negative Comments posted by users based on their opinions.

- o **Positive**: If the user comment contains at least one of the word which is listed in positive words, then that comment will be treated as positive comment.
- o **Negative**: If the user comment contains at least one of the word which is listed in negative words, then that comment will be treated as negative comment.

V. CONCLUSION

This study introduces a novel spam detection framework namely Plexus Spam based on a meta path concept as well as IEEE Transactions on Information Forensics and Security,Volume:12, Issue:7, Issue Date: 10 July, 2017 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 metapath 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, Plexus 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 behavioral 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.

For future work, meta path concept can be applied to other problems in this field. For example, similar framework can be used to find spammer communities. For finding community, reviews can be connected through group spammer features (such as the proposed feature in [29]) and reviews with highest similarity based on metapath concept are known as communities. In addition, utilizing the product features is an interesting future work on this study as we used features more related to spotting spammers and spam reviews. Moreover, while single networks has received considerable attention from various disciplines for over a decade, information diffusion and content sharing in multilayer networks is still a young research [37]. Addressing the problem of spam detection in such networks can be considered as a new research line in this field.

VI. 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.

[3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock.Finding deceptive opinion spam by any stretch of the imagination.In ACL, 2011.

[4] Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. In SIAM International Conference on Data Mining, 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.Exploitingburstiness in reviews for review spammer detection.In ICWSM, 2013.

[8] A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. Trueview: Harnessing the power of multiple review sites. In ACM WWW, 2015.

[9] 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.

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

[11] L. Akoglu, R. Chandy, and C. Faloutsos.Opinion fraud detection in online reviews bynetworkeffects.In ICWSM, 2013.

[12] R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review networksand metadata. In ACM KDD, 2015.

[13] S. Feng, R. Banerjee and Y. Choi.Syntacticstylometry for deception detection. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers; ACL, 2012.

[14] N. Jindal, B. Liu, and E.-P. Lim. Finding unusual review patterns using unexpected rules. In ACM CIKM, 2012.

[15] E.-P. Lim, V.-A.Nguyen, N. Jindal, B. Liu, and H. W. Lauw.Detecting product review spammers using rating behaviors.In ACM CIKM, 2010.

[16] A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh.Spotting opinion spammers using behavioral footprints.In ACM KDD, 2013.

[17] S. Xie, G. Wang, S. Lin, and P. S. Yu. Review spam detection via temporal pattern discovery. In ACM KDD, 2012.

[18] G. Wang, S. Xie, B. Liu, and P. S. Yu. Review graph based online store review spammer detection. IEEE ICDM, 2011.

[19] Y. Sun and J. Han. Mining Heterogeneous Information Networks; Principles and Methodologies, In ICCCE, 2012.

[20] A. Mukerjee, V. Venkataraman, B. Liu, and N. Glance. What Yelp Fake Review Filter Might Be Doing?, In ICWSM, 2013.