**AN EFFICIENT DETECTION METHOD OF EXTREMIST REVIEWER GROUPS IN ONLINE PRODUCT REVIEWS**

\*Dr. C. Premila Rosy1,

Head & Assistant Professor,

 PG & Dept., of CS,

Idhaya College for Women,

Kumbakonam,

Tamil Nadu, India

Ms. A. Fairosebanu2,

Assistant Professor of CS,

Idhaya College for Women,

Kumbakonam,

Tamil Nadu, India

Ms. B. Dhivyasri3,

II – M.Sc., (CS),

Idhaya College for Women,

Kumbakonam,

Tamil Nadu, India.

Ms. R. Joshi Anusha4,

II – M.Sc., (CS),

Idhaya College for Women,

Kumbakonam,

Tamil Nadu, India.

Ms. R. Swathi Priya5,

II – M.Sc., (CS),

Idhaya College for Women,

Kumbakonam,

Tamil Nadu, India.

Corresponding Author \*Dr. C. Premila Rosy – premilarosy78@gmail.com

**ABSTRACT**

Opinionated social media such as product reviews are now widely used by individuals and organizations for their decision making. However, due to the reason of profit or fame, people try to game the system by opinion spamming (e.g., writing fake reviews) to promote or demote some target products. When making purchasing decisions, customers increasingly rely on opinions posted on the Internet. Businesses therefore have an incentive to promote their own products or demote competitors’ products by creating positive or negative spam reviews on several platforms. Several researchers propose methods and tools to detect review spam automatically. In proposed system, collected reviews from the product review site and manually labeled a set of 923 candidate reviewer groups. The groups are extracted using frequent item set mining over brand similarities such that users are clustered together if they have mutually reviewed (products of) a lot of brands. To develop a feature-based supervised model to classify candidate groups as extremist entities. Then run multiple classifiers for the task of classifying a group based on the reviews written by the users of that group, to determine if the group shows signs of extremity. A 3-layer Preceptor based classifier turns out to be the best classifier. Further note that the proposed technique departs from the traditional supervised learning approach for spam detection because of the inherent nature of the problem which makes the classic supervised learning approach less effective. Experimental results show that the proposed method outperforms multiple strong baselines including the state-of-the-art supervised classification, regression, and learning to rank algorithms.

**Keywords:** Opinion Spam, Product Review, Rank Algorithm, Fake Reviews

**I. INTRODUCTION**

E-commerce is growing at an unprecedented rate all over the globe. With its growth, the impact of online reviews is increasing day by day. Reviews can influence people's purchasing decisions. Nowadays, reading product reviews before buying the product has become a habit, especially for potential customers. Customers post reviews about a product they purchase which may be positive or negative. Such reviews provide valuable feedback on these products, which may further be used by potential customers to find the opinions of existing users before deciding to purchase a product. If customers want to buy a product, they usually read reviews from some customers about the current product. If the reviews are mostly positive, there is a big chance to buy the product. Otherwise, if the reviews are mostly negative, customers tend to buy other products. While online reviews can be helpful, blind trust of these reviews is dangerous for both the seller and buyer. Most customers read online reviews before placing any online order. However, the reviews may be deceptive for extra profit or gain, thus any purchasing decision based on online reviews must be made carefully. To sell their products, companies often pursue customers to give desired reviews. There is a growing incentive for businesses to solicit and manufacture deceptive reviews, a.k.a. opinion spam- fictitious reviews that have been deliberately written to sound authentic and deceive the reader [1]. For example, Ott [2] has estimated that between 1% and 6% of positive hotel reviews appear to be deceptive, suggesting that some hotels may be posting fake positive reviews in order to hype their own offerings.

**II. OPINON SPAM DETECTION**

One key difficulty for opinion spam detection is that it is very hard to manually label fake reviews or reviewers for model building because it is almost impossible to recognize spam by just reading each individual review [3]. In this work, multiple experts were employed to create a labeled group opinion spammer dataset. This research makes the following main contributions:

1. It produces a labeled group spam dataset. To the best of our knowledge, this is the first such dataset. What was surprising and also encouraging to us was that unlike judging individual fake reviews or reviewers, judging fake reviewer groups were considerably easier due to the group context and their collective behaviours.

2. It proposes a novel relation-based approach to detecting spammer groups. With the labeled dataset, the traditional approach of supervised learning can be applied [3, 4, and 5]. However, we show that this approach can be inferior due to the inherent nature of our particular problem:

(i) Traditional learning assumes that individual instances are independent of one another. However, in our case, groups are clearly not independent of one another as different groups may share members. One consequence of this is that if a group *i* is found to be a spammer group, then the other groups that share members with group *i* are likely to be spammer groups too. The reverse may also hold.

(ii) It is hard for features used to represent each group in learning to consider each individual member’s behaviour on each individual product, i.e., a group can conceal a lot of internal details. This results in severe information loss, and consequently low accuracy. To exploit the relationships of groups, individual members, and products they reviewed, a novel relation-based approach is proposed, which call GSRank (Group Spam Rank), to rank candidate groups based on their likelihoods for being spam.

3. A comprehensive evaluation has been conducted to evaluate GSRank. Experimental results show that it outperforms many strong baselines including the state-of-the-art learning to rank, supervised classification and regression algorithms.

**III. RELATED WORK**

The problem of detecting review or opinion spam was introduced in [3], which used supervised learning to detect individual fake reviews. Duplicate and near duplicate reviews which are almost certainly fake reviews were used as positive training data. While [6] found different types of behaviour abnormalities of reviewers, [7] proposed a method based on unexpected class association rules and [5] employed standard word and part-of-speech (POS) n-gram features for supervised learning. [4] also used supervised learning with additional features. [8] used a graph-based method to find fake store reviewers. A distortion based method was proposed in [9]. None of them deal with group spam.

Jindal and Liu [10] made a pioneering effort to detect fake reviews. They introduced the problem of opinion spam and analyzed online reviews in three varieties - untruthful opinions, seller/brand only reviews (no product involved) and non-reviews using near-duplicate content as an indicator of fake reviews. Other studies dealing with the detection of review-level spam explored linguistic features of text [11], hand-made rules and combination of review and reviewer features. A probabilistic framework for the same has also been proposed. Ott et al. [11] synthesized fake hotel reviews using Amazon Mechanical Turk, whereas Jindal and Liu [10] worked on data scraped from Amazon and used content duplicity as ground-truth. Both of them worked with features at a review level. Jindal et al.andLi et al. mentioned the role of brands briefly, but the main focus was on fake reviews rather than extreme reviews.

The effect of fraud reviewer groups is more detrimental and subtle than individual fraud reviewers. The issue of manual labeling was addressed by considering a group of reviewers instead of individual reviews. Mukherjee et al. [12] showed that labeling a group of reviewers is considerably easier than labeling individual reviews. Other interesting studies that leverage metadata to characterize different entities in e-commerce sites can be observed, where products, reviews and users are classified simultaneously.



**Figure 1: Architecture**

**IV. EXPERIMENTAL EVALUATION**

**A. Algorithm:** GSRank

In this it includes state-of-art semi supervised supervised classification algorithm. By using GSRank, rank the groups based on their behavior to detect the spam. To label the review or calculate the spamicity of review this assigns the point as 1 for spam, 0.5 for borderline, and 0 for non-spam. To study the feasibility of labeling and the quality of judging uses the Fleiss’ multi-rater kappa method. Author considers some indicators for spamming activities as below:

(1) Group Time Window (GTW):

(2) Group Deviation (GD):

(3) Group Content Similarity:

(4) Group Member Content Similarity (GMCS): (5) Group Early Time Frame (GETF):

(6) Group Size Ratio (GSR):

(7) Group Size (GS):

(8) Group Support Count (GSUP):

All these features or group behaviors refer as f1 to f8.When group attains a feature f > 0, it is spam group. Also it is necessary to consider the individual member behavior as given below.

• Individual Rating Deviation (IRD):

• Individual Content Similarity (ICS):

• Individual Early Time Frame (IETF):

• Individual Member Coupling in a Group (IMC):

Trying to ensure a safe comparison, it compare our method named Evidential Group Spammers Detection (EGSD) with two previous works in which authors rely on the FIM (Frequent Itemset Mining) technique to generate the candidate groups and almost the same features used in our work. The first method introduced in Detecting Group Review Spam (DGRS) then computed the different indicators value and use the SVM rank algorithm to rank them, the other method proposed. Here focus on the Ranking Group Spam algorithm (GSRank) which relies on an iterative algorithm to effectively rank the group spammers.

**V. CLASSIFICATION EXPERIMENTS**

Group Spam Features (GSF) *f*1…*f*8: These are the proposed eight (8) group features presented here.

Individual Spammer Features (ISF): A set of features for detecting individual spammers was reported. Using these features, we represented each group with their average values of all the members of each group. We want to see whether such individual spammer features are also effective for groups.

Linguistic Features of reviews (LF): Word and POS (part-of-speech) n-gram features were shown to be effective for detecting individual fake reviews. Here, we want to see whether such features are also effective for spam groups. For each group, we merged its reviews into one document and represented it with these linguistic features.

**VI. CONCLUSION**

This paper proposed to detect group spammers in product reviews. The proposed method first used frequent item set mining to find a set of candidate groups, from which a labeled set of spammer groups was produced. Here found that although labeling individual fake reviews or reviewers is hard, labeling groups is considerably easier. Then proposed several behaviour features derived from collusion among fake reviewers. A novel relation-based model, called GSRank, was presented which can consider relationships among groups, individual reviewers, and products they reviewed to detect spammer groups. This model is very different from the traditional supervised learning approach to spam detection. Experimental results showed that GSRank significantly outperformed the state-of-the-art supervised classification, regression, and learning to rank algorithms. On further investigation, here came to know that this is indeed the case. One such confession by an anonymized top reviewer was seen in a blog post. Thus, it is to be expected that the online marketplace will be infested with manipulated extreme review instances.

**REFERENCES**

[1] B. Liu, “Sentiment analysis and opinion mining,” Synthesis lectures on human language technologies, vol. 5, no. 1, pp. 1–167, 2012.

[2] M. Ott, C. Cardie, and J. Hancock, “Estimating the prevalence of Deception in online review communities,” in Proceedings of the 21st International Conference on World Wide Web, pp. 201–210, ACM, 2012.

[3] Jindal, N. and Liu, B. Opinion spam and analysis. *WSDM*. 2008.

[4] Li, F., Huang, M., Yang, Y. and Zhu, X. Learning to identify review Spam. *IJCAI*. 2011.

[5] Ott, M., Choi, Y., Cardie, C. Hancock, J. Finding Deceptive Opinion Spam by Any Stretch of the Imagination. *ACL*. 2011.

[6] Lim, E. Nguyen, V. A., Jindal, N., Liu, B., and Lauw, H. Detecting Product Review Spammers Using Rating Behavior. *CIKM*. 2010.

[7] Jindal, N., Liu, B. and Lim, E. P. Finding Unusual Review Patterns Using Unexpected Rules. *CIKM*. 2010.

[8] Wang, G., Xie, S., Liu, B., and Yu, P. S. Review Graph based Online Store Review Spammer Detection. *ICDM*. 2011.

[9] Wu, G., Greene, D., Smyth, B. and Cunningham, P. 2010. Distortion as a validation criterion in the identification of suspicious reviews.Technical report, UCD-CSI-2010-04, University College Dublin.

[10] N. Jindal and B. Liu, “Opinion spam and analysis,” in WSDM, 2008,pp. 219–230.

[11] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, “Finding deceptive opinion spam by any stretch of the imagination,” in ACL, 2011, pp.309–319.

[12] A. Mukherjee, B. Liu, and N. Glance, “Spotting fake reviewer groups in consumer reviews,” in WWW, 2012, pp. 191–200.