ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 2, February 2020

An Effective Filtering Based Fraud and Fake Reviews Detection in E-Commerce Platform

R.Sowmiya[1], R.Ashwinidevi[2], C.Esaivani[2], S.Anusuya[2]

Assistant Professor, Department of Computer Science and Engineering, Mahendra Institute of Technology, Namakkal, India[1]

Department of Computer Science and Engineering, Mahendra Institute of Technology, Namakkal, India[2]

ABSTRACT: Now a day's ranking fraud in mobile App market became more popular in the market in order to display their apps in the popularity list and to boost their sales .Therefore, there will be increase in ranking fraud in the upcoming time, as the number of Apps developers and applications will likely grow very significantly. Many traditional methods of fraud analysis have been used to detect fraud. But these methods are complex and time consuming. However there is more need to adopt some better techniques which can ensure the ranking fraud detection efficiently by data mining analysis. This paper explores the data mining methods to identify the fraud by using Rank Aggregation algorithm. Further three types of proofs are studied they are ranking, rating and review proofs. And an aggregation method is used to aggregate all the proofs and will produce an optimized report for fraud detection. We demonstrate the effectiveness of our framework on synthetic and real datasets; where FRAUDEAGLE successfully reveals fraud-bots in a large online app review database

I. INTRODUCTION

The Web has greatly enhanced the way people perform certain activities (e.g. shopping), find information, and interact with others. Today many people read/write reviews on merchant sites, blogs, forums, and social media before/after they purchase products or services. Examples include restaurant reviews on Yelp, product reviews on Amazon, hotel reviews on TripAdvisor, and many others. Such user-generated content contains rich information about user experiences and opinions, which allow future potential customers to make better decisions about spending their money, and also help merchants improve their products, services, and marketing. Since online reviews can directly influence customer purchase decisions, they are crucial to the success of businesses. While positive reviews with high ratings can yield financial gains, negative reviews can damage reputation and cause monetary loss. This effect is magnified as the information spreads through the Web (Hitlin 2003; Mendoza, Poblete, and Castillo 2010). As a result, online review systems are attractive targets for opinion fraud. Opinion fraud involves reviewers (often paid) writing bogus reviews (Kost May 2012; Streitfeld August 2011). These spam reviews come in two flavors: defaming-spam which untruthfully vilifies, or hypespam that deceitfully promotes the target product.

The opinion fraud detection problem is to spot the fake reviews in online sites, given all the reviews on the site, and for each review, its text, its author, the product it was written for, timestamp of posting, and its star-rating. Typically no user profile information is available (or is self-declared and cannot be trusted), while more side information for products (e.g. price, brand), and for reviews (e.g. number of (helpful) feedbacks) could be available depending on the site.

Detecting opinion fraud, as defined above, is a non-trivial and challenging problem. Fake reviews are often written by experienced professionals who are paid to write high quality, believable reviews. As a result, it is difficult for an average potential customer to differentiate bogus reviews from truthful ones, just by looking at individual reviews

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 2, February 2020

text(Ott et al. 2011). As such, manual labeling of reviews is hard and ground truth information is often unavailable, which makes training supervised models less attractive for this problem.

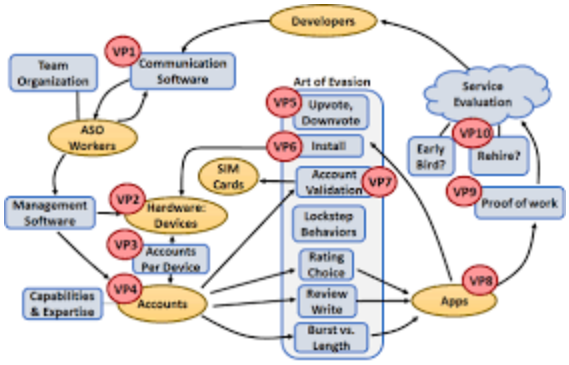

Fig 1: System Flow

Summary of previous work. Previous attempts at solving the problem use several heuristics, such as duplicated reviews (Jindal and Liu 2008), or acquire bogus reviews from non-experts (Ott et al. 2011), to generate pseudo-ground truth, or a reference dataset. This data is then used for learning classification models together with carefully engineered features. One downside of such techniques is that they do not generalize: one needs to collect new data and train a new model for review data from a different domain, e.g., hotel vs. restaurant reviews. Moreover feature selection becomes a tedious sub-problem, as datasets from different domains might exhibit different characteristics. Other feature-based proposals include (Lim et al. 2010; Mukherjee, Liu, and Glance 2012).

A large body of work on fraud detection relies on review text information (Jindal and Liu 2008; Ott et al. 2011; Feng, Banerjee, and Choi 2012) or behavioral evidence (Lim et al. 2010; Xie et al. 2012; Feng et al. 2012), and ignore the connectivity structure of review data. On the other hand, the network of reviewers and products contains rich information that implicitly represents correlations among these entities. The review network is also invaluable for detecting teams of fraudsters that operate collaboratively on targeted products. Our contributions. In this work we propose an unsuper- vised, general, and network-based framework, FRAUDEAGLE, to tackle the opinion fraud detection problem in online review data. The review network successfully captures the correlations of labels among users and products, e.g. fraudsters are mostly linked to good (bad) products with negative (positive) fake reviews, and vice versa for honest users. As such, the network edges are signed by sentiment. We build an iterative, propagation-based algorithm that exploits the network structure and the long-range correlations to infer the class labels of users, products, and reviews. A second step involves analysis and summarization of results. For generality, we do not use review text information, but only the positive or negative sentiment of the reviews. As such, our method can be applied to any type of review data and is complementary to existing approaches

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 2, February 2020

Fig 2: Work Flow

Availability of a dataset is the key starting point of any spam review detection research. The key issue in the spam review detection problem is the availability of the labeled dataset. It has been observed from existing studies that only one labeled hotel review dataset is available [15], but it has only review text and no other features are available, such as review posting time and review rating. Researchers need to have access to the labeled dataset to train a classifier so that it may classify an unknown review as spam or not-spam. The alternative approach is to use an artificially created review dataset by using synthetic review spamming [16]. Ott et al. [15] have used Amazon Mechanical Turk (www.mturk.com) (AMT) for collecting labeled datasets through online workers (called Tuckers) to write fake hotel reviews (by giving one dollar for each review) in order to represent a few hotels in a positive light and some negatively. According to Mukherjee et al. [4], the process of labeling has not provided improved accuracy for spam review detection on real-life datasets. Table 6 lists review datasets used by different researchers and show total reviews, number of reviewers, and number of products for each dataset. It is observed by the review of the literature that all review datasets are not publicly available, and many times researchers use crawlers to gather required data. It has also been observed that most of the researchers used Amazon.com e-commerce website datasets in their works, as it is the biggest e-commerce platform to have product reviews, and the second largest review dataset is available from tripadvisor.com, which is an online hotel booking website. Additionally, the researchers working in the spam review detection area use these datasets provided by such websites.

II. LITERATURE REVIEW

The systematic literature review provides the answers to specific research questions, whereas the general survey paper gives a broad idea about the domain. This SLR is performed using the guidelines provided in [10] and the selection of the primary studies is performed according to [11]. The key objective of this research is to identify the best available feature extraction techniques, and present different existing models for spam review detection and available parameters to analyze these models. The process of SLR helps to determine different studies available in the domain of spam review detection and answer different research questions. The distinct phases of the systematic literature review.

Much of the previous work in opinion fraud focuses on review text content, behavioral analysis, and supervised methods. (Jindal and Liu 2008) identified opinion spam by detecting exact text duplicates in an Amazon.com dataset, while (Ott et al. 2011) crowd-sourced deceptive reviews in order to create a highly accurate classifier based on ngrams. Several studies tried to engineer better features to improve classifier performance. (Li et al. 2011) uses sentiment scores, product brand, and reviewer profile attributes to train classifiers. Other work has computed scores based on behavioral heuristics, such as rating deviation by (Lim et al. 2010), and frequent itemset mining to find fraudulent reviewer groups by (Mukherjee, Liu, and Glance 2012). Unfortunately, these methods are not generalizable: the models need re-training to account for differences between problem domains, such as book reviews versus movie reviews.

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 2, February 2020

Moreover, the features might not be consistent even for datasets within the same domain, depending on the dataset source. Consequently, feature extraction becomes a time-consuming yet pivotal sub-problem with attributes varying across domains. Another group of work mines behavioral patterns in review data. (Jindal, Liu, and Lim 2010) finds unexpected rules to highlight anomalies, and (Xie et al. 2012; Feng et al. 2012; Feng, Banerjee, and Choi 2012) respectively study temporal review behavior, rating distributions, and syntactic stylometry. On the other hand, methods that account for the network of reviewers, reviews, and products can more elegantly encapsulate structural signals that go beyond the review content and simple heuristics, thus generalizing across domains. (Wang et al. 2011) proposed the first (and to the authors' knowledge the only) review graph-based method and a simple yet effective algorithm to compute scores for each node. We propose an even more flexible method that exploits the network effects; being able to incorporate side information, is based on the rigorous theoretical foundation of belief propagation, and is linearly scalable. Network effects have been exploited in securities fraud (Neville et al. 2005), accounting fraud (McGlohon et al. 2009), and auction fraud (Pandit et al. 2007) detection. However, none of these proposals are applicable to opinion fraud detection, as their problem domains do not involve ratings and sentiment spam. Related to ratings on products, work on recommender systems (Koren 2009; Menon and Elkan 2011), aim for best prediction of future user ratings, but do not address the fraud problem.

III. PROBLEM DESCRIPTION AND TOY EXAMPLE

Simply put, we consider the problem of spotting fraudulent reviewers, and consequently spotting fake reviews in online review datasets. The online review datasets mainly consist of a set of users (also called customers, reviewers), a set of products (e.g., hotels, restaurants, etc.), and the reviews. Each review is written from a particular user to a particular product, and contains a star-rating, often an integer from 1 to 5. As such, a review dataset can be represented as a bipartite network. In this network, user nodes are connected to product nodes, in which the links represent the "reviewed" relationships and each link is associated with the review rating. The objects in the review network, i.e. the users, products, and reviews, can be grouped into certain classes. In this paper, we consider two classes for each object type: products are either good or bad quality, users are either honest or fraud, and finally reviews are either real or fake. Intuitively, a product is good (bad) if it most of the time receives many positive (negative) reviews from honest users. Similarly, a user is honest (fraud) if s/he mostly writes positive (negative) reviews to good products, and negative (positive) reviews to bad products. In other words, a user is fraud if s/he is trying to promote a set of target bad products (hypespam), and/or damage the reputation of a set of target good products (defaming-spam). All the reviews of honest users can safely be regarded as real. In an ideal setting, all the reviews of the fraud users can also be thought of as fake. However, in reality fraudsters could also write realistic reviews, trying to hide their otherwise fraudulent activity. All in all, notice that the class labels of the interconnected objects in the review data are strongly correlated as they are described in terms of one another. As a result, it is natural to think of a review dataset as a network in order to attack the opinion fraud problem coherently. Given a user-product review network as described above, the task is to assign each object with a class label that best describes the observed data, i.e. the network structure and ratings. In this task, we need to infer from review text whether it is of positive or negative sentiment. While there exists prior work on sentiment classification using text information (Pang, Lee, and Vaithyanathan 2002), we take the review rating as a highly correlated indicator for sentiment. More specifically, we consider a signed network in which each network link (i.e. review) is marked as positive if its rating is above a threshold, and as negative otherwise.

IV. PROPOSED TAXONOMY

Based on the analysis of primary studies, this work proposes a new taxonomy that may be used by other researchers to classify existing approaches and to figure out the most appropriate technique to solve a spam review detection problem. It is observed that there are 2 major approaches to handle the spam review detection problem: Machine learning and Lexicon-based approaches. Further, these main approaches have been classified into different methods with associated classes to resolve the problem of spam review detection. The proposed taxonomy is presented in Figure 4. To the best

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 2, February 2020

of the researchers' knowledge, this is the first attempt to compile all the different approaches pertaining to spam review detection.

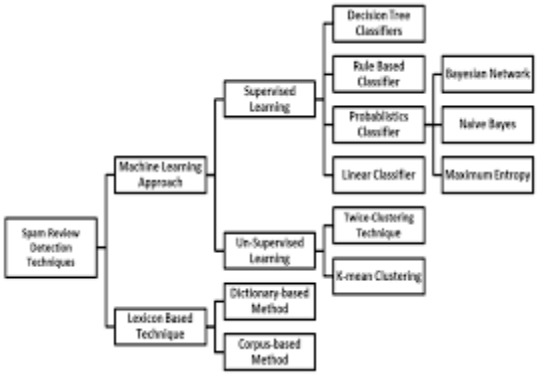

Fig 3: Proposed Taxonomy Review

V. SUPERVISED LEARNING

Supervised learning approaches used for spam review detection are commonly based on the classification methods [15]. In this learning technique, two datasets are required: training data and test data. Training data is utilized to train the classifier and afterwards test data is utilized to evaluate the performance of a classifier [28,46]. Methods such as Support Vector Machine (SVM) and Naïve Bayes (NB) have already shown great success in opinion mining. Researchers usually start by gathering and crawling the dataset. The next step is to prepare and pre-process the dataset according to the domain. Once the dataset is prepared then features are extracted from the dataset by using the feature engineering approach. The next step is to train the classifier by using training data.

VI. DECISION TREE CLASSIFIER

Decision tree (DT) classifiers give a hierarchical decomposition of the training data space and are used to learn the rules to identify the authenticity of the review [47,48]. A tree is formed by using different features and their values. Information gain is calculated by using a list of features. The feature that has maximum information gain is used as the root node of the decision tree. The interior nodes of the decision tree are labeled with unique features and these features have low information gain as compared to the root node. This procedure is repeated until all reviews are classified as spam or not-spam reviews. In the study by Jotheeswaran et al. [49], the IMDb movie reviews dataset is used, and inverse document frequency method is used to extract unique features and decision tree induction selects relevant features. They claimed to have correctly classified 75% of the reviews as spam or not-spam.

VII. CONCLUSION

we have addressed fake review detection in the consumer electronics domain. We have proposed a feature framework oriented to analysis of social sites, and we have also developed a dataset which is made available2 for future research. Our framework is composed of two main types of features: review centric and user centric. Review centric features are only related to the text of the review. On the other hand, user centric features show how the user behaves within the site, and are subdivided into four groups: personal, social, reviewing activity and trust. As shown in the article, detecting fake reviews by just reading the reviews is a challenging task for both humans and computers, since textual reviews do not usually provide fake signals to be detected. In contrast, features related to the user have been shown to be more effective. The most relevant ones are those coming from the reviewing activity. This is not surprising, since

ISSN(Online): 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 2, February 2020

they can provide signals of abnormal reviews. The rest of features have also proven to have discrimination capability and every combination of feature types have improved the overall accuracy.

REFERENCES

1. Xue, H.; Li, F.; Seo, H.; Pluretti, R. Trust-aware review spam detection. In Proceedings of the 2015 IEEE Trustcom/BigDataSE/ISPA, Helsinki, Finland, 20–22 August 2015; Volume 1, pp. 726–733.
2. Lau, R.Y.; Liao, S.Y.; Kwok, R.C.; Xu, K.; Xia, Y.; Li, Y. Text mining and Probabilistic language modeling for online review spam detection. ACM Trans. Manag. 2011, 2, 25.
3. Serrano-Guerrero, J.; Olivas, J.A.; Romero, F.P.; Herrera-Viedma, E. Sentiment analysis: A review and comparative analysis of web services. Inf. Sci. 2015, 311, 18–38.
4. Mukherjee, A.; Venkataraman, V.; Liu, B.; Glance, N.S. What Yelp fake review filter might be doing? In Proceedings of the International Conference on Web and Social Media, Cambridge, MA, USA, 8–11 July 2013; pp. 409–418.
5. Rashid, A.; Anwer, N.; Iqbal, M.; Sher, M. Areas, Techniques, Challenges of Opinion Mining. Int. J. Comput. Sci. 2013, 10, 18–31.
6. Mukherjee, A.; Venkataraman, V.; Liu, B.; Glance, N.S. A Survey on Detection of Shill Reviews by Measuring its Linguistic Features. Int. J. Emerg. Trends Technol. Comput. Sci. (Ijettcs) 2014, 3, 269–272. 7. Spirin, N.; Han, J. Survey on web spam detection: Principles and algorithms. ACM Sigkdd Explor. Newsl. 2011, 13, 50–64.
8. Chakraborty, M.; Pal, S.; Pramanik, R.; Chowdary, C.R. Recent developments in social spam detection and combating techniques: A survey. Inf. Process. Manag. 2016, 52, 1053–1073.
9. Peng, J.; Choo, K.K.; Ashman, H. User profiling in intrusion detection: A review. J. Netw. Comput. Appl. 2016, 72, 14–27.
10. Keele, S. Guidelines for Performing Systematic Literature Reviews in Software Engineering; Ver. 2.3 EBSE Technical Report: Software Engineering Group, School of Computer Science and Mathematics Keele University, UK and Department of Computer Science University of Durham, UK: 2007; pp. 1–57.
11. Kitchenham, B. Procedures for Undertaking Systematic Reviews; Joint technical report; Computer Science Department, Keele University (TR/SE-0401) and National ICT,SydneyAustralia Ltd.: 2004; Volume 51, pp. 7–15. Appl. Sci. 2019, 9, 987 23 of 26
12. Wohlin, C. Guidelines for snowballing in systematic literature studies and replication in software engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, London, UK, 13–14 May 2014; p. 38