International Journal of Research and Development in Applied Science and Engineering (IJRDASE)
ISSN: 2454-6844

Network-Based Spam Detection in Online Reviews: A Novel Graph-Theoretic Framework for Social Media Integrity

Km Pallavi Singh[1], Rohitashwa Pandey[2]
Dept. of Computer Science and Engineering,
Bansal Institute of Engineering, (AKTU), Lucknow, India

Abstract— **In the digital age, online reviews significantly influence consumer decisions and business reputations. However, the proliferation of spam reviews, often generated by coordinated malicious actors, poses a serious threat to the reliability of such platforms. This paper presents a novel graph-theoretic framework for network-based spam detection in online social media reviews. By representing user-review interactions as dynamic graphs, the proposed approach captures underlying behavioral patterns and relational anomalies indicative of spam activity. Central to this method is the identification of structural inconsistencies, such as dense subgraphs formed by collusive user groups and anomalous propagation patterns across the network. We incorporate key graph metrics—such as centrality, modularity, and clustering coefficients—to distinguish spammers from legitimate users with high accuracy. Our framework was evaluated on multiple real-world datasets, demonstrating superior performance in spam detection compared to traditional machine learning-based classifiers. This work contributes to enhancing trust and transparency in online platforms by offering a robust, scalable, and interpretable solution for preserving the integrity of user-generated content.**

Keywords: Network-based spam detection, online reviews, graph theory, social media integrity, review spam, collusion detection, user behavior analysis, graph metrics, community detection, fake review identification.

1. INTRODUCTION

The dissemination of information through online social media portals is a significant factor that producers use in their advertising campaigns and customers use to make product and service selections. People these days heavily rely on written reviews to make decisions, with positive and negative reviews encouraging and discouraging them in their product and service selections. Written reviews also aid service providers in raising the level of quality of their goods and services. As a result, these reviews are now a big part of a company's success. While good reviews can help a business, bad reviews can hurt its credibility and cost the company money. Spammers will take advantage of the fact that anyone can post comments pretending to be reviews, creating a tempting opportunity for them to write fake reviews with the intention of misleading users. The sharing feature of social media and the spread of these false reviews online add to their number. When making or canceling a purchase, consumers are increasingly relying on user-generated online reviews. As a result, it appears that businesses and the general public are increasingly concerned about the possibility of publishing deceptive opinion spam|citations reviews that have been purposefully written to sound authentic and deceive the reader. Although this may come as a surprise, very little is known about the actual rate of deception in online review communities and even less about the factors that may contribute to it. On the one hand, the pressure to portray businesses, products, and services in a positive light and the relative ease with which reviews can be written could lead one to believe that most online reviews are fake. On the other hand, it could be argued that review websites cannot provide any value unless there is a low rate of deception. In the context of online reviews, the detection of spam has been the primary focus of research. To identify duplicate opinions, Jindal and Liu, for instance, train models with features derived from the reviewer, product, and review text.

Using a standard statistical test, manually compare the psychologically relevant linguistic differences between 40 honest and 42 false hotel reviews. These strategies are useful, but they do not address the prevalence of deception in online reviews. In point of fact, empirical and academic studies of the extent of deceptive opinion spam remain scarce. One reason is that it's hard to find trustworthy gold-standard annotations for reviews, such as labels that indicate whether a review is honest (real) or deceptive (fake). One choice for creating highest quality level marks, for instance, is depend on the decisions of human annotators. However, recent research demonstrates that human readers have difficulty identifying deceptive opinion spam; this is particularly the situation while considering the over believing nature of most human adjudicators, a peculiarity alluded to in the mental double dealing writing as a reality predisposition. Recent large meta-analyses demonstrating the inaccuracy of human judgments of deception, with accuracy rates typically close to chance, support the difficulty of identifying fake reviews.

Even though self-reports are difficult and costly to obtain, especially in large-scale settings like the internet, it is not surprising that research on estimating the prevalence of deception has generally relied on self-report methods because humans have difficulty recognizing deceptive messages from cues alone. In addition, self-report methods like diaries and large-scale surveys have a number of methodological issues, such as self-deception and social desirability bias. In addition, revealing one's own deception in online reviews is severely discouraged by the possibility of permanent exclusion from a review portal or damage to a company's reputation. According to signaling theory, each review in our situation is a sign of the product's true, unknowable quality; As a result, the objective of consumer reviews is to reduce the inherent information gap between consumers and manufacturers. In a nutshell, the prevalence of deception ought to be proportional to the costs and benefits of fabricating a review, in accordance with a signaling theory approach. We hypothesize that review communities with low signaling costs, like those that make it simple to post a review, and large benefits, like sites with a lot of traffic, will have more deceptive opinion spam than communities with higher signaling costs, like those that make additional requirements for posting reviews, and lower benefits, like low site traffic. It is now common knowledge that user-generated content contains useful data that can be used for a variety of purposes. We focus on product reviews from customers in this paper. We focus primarily on reviewing opinion spam. User feedback on goods and services is abundant in reviews. Before making a purchase decision, potential customers use them to learn what other people think of a product.

They are also used by product manufacturers to identify product problems and/or to find marketing intelligence information about their competitors. In the past few years, there was a growing interest in mining opinions in reviews from both academia and industry. However, the existing work has been mainly focused on extracting and summarizing opinions from reviews using natural language processing and data mining techniques. Little is known about the characteristics of reviews and behaviors of reviewers. There is also no reported study on the trustworthiness of opinions in reviews. Due to the fact that there is no quality control, anyone can write anything on the Web. This results in many low quality reviews, and worse still *review spam*. Review spam is similar to Web page spam. In the context of Web search, due to the economic and/or publicity value of the rank position of a page returned by a search engine, Web page spam is widespread. Web page spam refers to the use of “illegitimate means” to boost the rank positions of some target pages in search engines. In the context of reviews, the problem is similar, but also quite different. It is now very common for people to read opinions on the Web for many purposes. For example, if one wants to buy a product and sees that the reviews of the product are mostly positive, one is very likely to buy it. If the reviews are mostly negative, one is very likely to choose another product. Positive opinions can result in significant financial gains and/or fames for organizations and individuals. This gives good incentives for review/opinion spam.

2. RELATED WORK

The On the basis of extensive literature survey related to Spam Detection in Online Social Media Using a Network-based Framework for Reviews has been taken into consideration in this section.

E. D. Wahyuni (2016) suggested that the rapid growth of the Internet influenced many of our daily activities. One of the very rapid growth areas is ecommerce. Generally e-commerce provides facility for customers to write reviews related with its service. The existence of these reviews can be used as a source of information. For examples, companies can use it to make design decisions of their products or services, while potential customers can use it to decide either to buy or to use a product. Unfortunately, the importance of the review is misused by certain parties who tried to create fake reviews, both aimed at raising the popularity or to discredit the product. This research aims to detect fake reviews for a product by using the text and rating property from a review. In short, the proposed system (ICF++) will measure the honesty value of a review, the trustiness value of the reviewers and the reliability value of a product. The honesty value of a review will be measured by utilizing the text mining and opinion mining techniques. The result from the experiment shows that the proposed system has a better accuracy compared with the result from iterative computation framework (ICF) method.

M. Crawford (2016) suggested that online reviews are quickly becoming one of the most important sources of information for consumers on various products and services. With their increased importance, there exists an increased opportunity for spammers or unethical business owners to create false reviews in order to artificially promote their goods and services or smear those of their competitors. In response to this growing problem, there have been many studies on the most effective ways of detecting review spam using various machine learning algorithms. One common thread in most of these studies is the conversion of reviews to word vectors, which can potentially result in hundreds of thousands of features. However, there has been little study on reducing the feature subset size to a manageable number or how best to do so. In this paper, we consider two distinct methods of reducing feature subset size in the review spam domain. The methods include filter-based feature rankers and word-frequency based feature selection. We show that there is not a one size fits all approach to feature selection, and the best way to reduce the feature subset size is dependent upon both the classifier being used and the feature subset size desired. It was also observed that the feature subset size had significant influence on which feature selection method is used.

M. Luca and G. Zervas (2016) suggested that Consumer reviews are now part of everyday decision-making. Yet, the credibility of these reviews is fundamentally undermined when businesses commit review fraud, creating fake reviews for themselves or their competitors. We investigate the economic incentives to commit review fraud on the popular review platform Yelp, using two complementary approaches and datasets. We begin by analysing restaurant reviews that are identified by Yelp’s filtering algorithm as suspicious, or fake – and treat these as a proxy for review fraud (an assumption we provide evidence for). We present four main findings. First, roughly 16% of restaurant reviews on Yelp are filtered. These reviews tend to be more extreme (favorable or unfavorable) than other reviews, and the prevalence of suspicious reviews has grown significantly over time. Second, a restaurant is more likely to commit review fraud when its reputation is weak, i.e., when it has few reviews, or it has recently received bad reviews. Third, chain restaurants – which benefit less from Yelp – are also less likely to commit review fraud. Fourth, when restaurants face increased competition, they become more likely to receive unfavorable fake reviews. Using a separate dataset, we analyze businesses that were caught soliciting fake reviews through a sting conducted by Yelp. These data support our main results, and shed further light on the economic incentives behind a business’s decision to leave fake reviews.

A. j. Minnich (2015) suggested that online reviews on products and services can be very useful for customers, but they need to be protected from manipulation. So far, most studies have focused on analyzing online reviews from a single hosting site. How could one leverage information from multiple review hosting sites? This is the key question in our work. In response, we develop a systematic methodology to merge, compare, and evaluate reviews from multiple hosting sites. We focus on hotel reviews and use more than 15 million reviews from more than 3.5 million users spanning three prominent travel sites. Our work consists of three thrusts: (a) we develop novel features capable of identifying cross-site discrepancies effectively, (b) we conduct arguably the first extensive study of cross-site variations using real data, and develop a hotel identity-matching method with 93% accuracy, (c) we introduce the True View score, as a proof of concept that cross-site analysis can better inform the end user. Our results show that: (1) we detect 7 times more suspicious hotels by using multiple sites compared to using the three sites in isolation, and (2) we find that 20% of all hotels appearing in all three sites seem to have low trustworthiness score. Our work is an early effort that explores the advantages and the challenges in using multiple reviewing sites towards more informed decision making.

R. Shebuti (2015) suggested that online reviews capture the testimonials of "real" people and help shape the decisions of other consumers. Due to the financial gains associated with positive reviews, however, opinion spam has become a widespread problem, with often paid spam reviewers writing fake reviews to unjustly promote or demote certain products or businesses. Existing approaches to opinion spam have successfully but separately utilized linguistic clues of deception, behavioral footprints, or relational ties between agents in a review system. In this work, we propose a new holistic approach called Spangle that utilizes clues from all metadata (text, timestamp, rating) as well as relational data (network), and harness them collectively under a unified framework to spot suspicious users and reviews, as well as products targeted by spam. Moreover, our method can evidently and seamlessly integrate semi-supervision, i.e., a (small) set of labels if available, without requiring any training or changes in its underlying algorithm. We demonstrate the electiveness and scalability of Spangle on three real-world review datasets from Yelp.com with filtered (spam) and recommended (non spam) reviews, where it significantly outperforms several baselines and state-of-the-art methods. To the best of our knowledge, this is the largest scale quantitative evaluation performed to date for the opinion spam problem.

B. Viswanath (2014) suggested that Users increasingly rely on crowd sourced information, such as reviews on Yelp and Amazon, and liked posts and ads on Facebook. This has led to a market for black hat promotion techniques via fake (e.g., Sybil) and compromised accounts, and collusion networks. Existing approaches to detect such behavior relies mostly on supervised (or semi-supervised) learning over known (or hypothesized) attacks. They are unable to detect attacks missed by the operator while labeling, or when the attacker changes strategy. We propose using unsupervised anomaly detection techniques over user behavior to distinguish potentially bad behavior from normal behavior. We present a technique based on Principal Component Analysis (PCA) that models the behavior of normal users accurately and identifies significant deviations from it as anomalous. We experimentally validate that normal user behavior (e.g., categories of Facebook pages liked by a user, rate of like activity, etc.) is contained within a low-dimensional subspace amenable to the PCA technique. We demonstrate the practicality and effectiveness of our approach using extensive ground-truth data from Facebook: we successfully detect diverse attacker strategies—fake, compromised, and colluding Facebook identities—with no a priori labeling while maintaining low false-positive rates. Finally, we apply our approach to detect click-spam in Facebook ads and find that a surprisingly large fraction of clicks are from anomalous users.

Ch. Xu and J. Zhang (2014) suggested that Spam campaigns spotted in popular product review websites (e.g., amazon.com) have attracted mounting attention from both industry and academia, where a group of online posters are hired to collaboratively craft deceptive reviews for some target products. The goal is to manipulate perceived reputations of the targets for their best interests. Many efforts have been made to detect such colluders by extracting point wise features from individual reviewers/reviewer-groups, however, pairwise features which can potentially capture the underlying correlations among colluders are either ignored or just explored insufficiently in the literature. We observed that pairwise features can be more robust to model the relationships among colluders since them, as the ingredients of spam campaigns, are correlated in nature. In his paper, we explore multiple heterogeneous pairwise features in virtue of some collusion signals found in reviewers' rating behaviors and linguistic patterns. In addition, an unsupervised and intuitive colluder detecting framework has been proposed which can benefit from these pairwise features. Extensive experiments on real dataset show the effectiveness of our method and satisfactory superiority over several competitors.

H. Li (2014) suggested that online reviews have become an increasingly important resource for decision making and product designing. But reviews systems are often targeted by opinion spamming. Although fake review detection has been studied by researchers for years using supervised learning, ground truth of large scale datasets is still unavailable and most of existing approaches of supervised learning are based on pseudo fake reviews rather than real fake reviews. Working with Dianping1, the largest Chinese review hosting site, we present the first reported work on fake review detection in Chinese with filtered reviews from Damping's fake review detection system. Damping's algorithm has a very high precision, but the recall is hard to know. This means that all fake reviews detected by the system are almost certainly fake but the remaining reviews (unknown set) may not be all genuine. Since the unknown set may contain many fake reviews, it is more appropriate to treat it as an unlabeled set. This calls for the model of learning from positive and unlabeled examples (PU learning). By leveraging the intricate dependencies among reviews, users and IP addresses, we first propose a collective classification algorithm called Multi-typed Heterogeneous Collective Classification (MHCC) and then extend it to Collective Positive and Unlabeled learning (CPU). Our experiments are conducted on real-life reviews of 500 restaurants in Shanghai, China. Results show that our proposed models can markedly improve the F1 scores of strong baselines in both PU and non-PU learning settings. Since our models only use language independent features, they can be easily generalized to other languages.

G. Fei (2013) suggested that online product reviews have become an important source of user opinions. Due to profit or fame, imposters have been writing deceptive or fake reviews to promote and/or to demote some target products or services. Such imposters are called review spammers. In the past few years, several approaches have been proposed to deal with the problem. In this work, we take a different approach, which exploits the burstiness nature of reviews to identify review spammers. Bursts of reviews can be either due to sudden popularity of products or spam attacks. Reviewers and reviews appearing in a burst are often related in the sense that spammers tend to work with other spammers and genuine reviewers tend to appear together with other genuine reviewers. This paves the way for us to build a network of reviewers appearing in different bursts. We then model reviewers and their concurrence in bursts as a Markov Random Field (MRF), and employ the Loopy Belief Propagation (LBP) method to infer whether a reviewer is a spammer or not in the graph. We also propose several features and employ feature induced message passing in the LBP framework for network inference. We further propose a novel evaluation method to evaluate the detected spammers automatically using supervised classification of their reviews. Additionally, we employ domain experts to perform a human evaluation of the identified spammers and non-spammers. Both the classification result and human evaluation result show that the proposed method outperforms strong baselines, which demonstrate the effectiveness of the method.

M. Ott (2012) suggested that Consumers' purchase decisions are increasingly inuenced by user-generated online reviews. Accordingly, there has been growing concern about the potential for posting deceptive opinion spam| citations reviews that have been deliberately written to sound authentic, to deceive the reader. But while this practice has received considerable public attention and concern, relatively little is known about the actual prevalence, or rate, of deception in online review communities, and less still about the factors that inuence it. We propose a generative model of deception which, in conjunction with a deception classifier, we use to explore the prevalence of deception in six popular online review communities: Expedia, Hotels.com, Orbits, Priceline, Trip Advisor, and Yelp. We additionally propose a theoretical model of online reviews based on economic signaling theory, in which consumer reviews diminish the inherent information asymmetry between consumers and producers, by acting as a signal to a product's true, unknown quality. We find that deceptive opinion spam is a growing problem overall, but with different growth rates across communities. These rates, we argue, are driven by the different signaling costs associated with deception for each review community, e.g., posting requirements. When measures are taken to increase signaling cost, e.g., filtering reviews written by first-time reviewers, deception prevalence is effectively reduced.

F. Li (2011) suggested that in the past few years, sentiment analysis and opinion mining becomes a popular and important task. These studies all assume that their opinion resources are real and trustful. However, they may encounter the faked opinion or opinion spam problem. In this paper, we study this issue in the context of our product review mining system. On product review site, people may write faked reviews, called review spam, to promote their products, or defame their competitors' products. It is important to identify and filter out the review spam. Previous work only focuses on some heuristic rules, such as helpfulness voting, or rating deviation, which limits the performance of this task. In this paper, we exploit machine learning methods to identify review spam. Toward the end, we manually build a spam collection from our crawled reviews.

3. TECHNIQUE AND ALGORITHMS

- **Sentiment Analysis**

In this algorithm, preprocessed tweets are brought from the database one by one. In the first place we require check one by one watchword whether that catchphrase is thing are not, if thing we will expel it from the specific review. After that the rest of the watchwords checked with assessment compose, regardless of whether that catchphrases are certain opinion or negative conclusion or impartial feeling. The rest of the watchwords in the tweet which does not has a place with any of the supposition will be relegated transitory conclusion in light of the more check of positive, negative and impartial. In the second cycle if the reaming word crosses the limit of positive, negative or nonpartisan, that watchword forever included as development in the lexicon.

Cosine Similarity calculation

Cosine similarity is a measure of similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them. The cosine of 0° is 1, and it is less than 1 for any angle in the interval $(0,\pi]$ radians. It is thus a judgment of orientation and not magnitude: two vectors with the same orientation have a cosine similarity of 1, two vectors oriented at 90° relative to each other have a similarity of 0, and two vectors diametrically opposed have a similarity of -1, independent of their magnitude. The cosine similarity is particularly used in positive space, where the outcome is neatly bounded .

Algorithm Step in Cosine Similarity

Step 1: Data Preparation

As with the k-means section, we will limit the number of attributes in the data set to A3 and A4 (petal length and petal width) using the Select Attribute operator, so that we can visualize the cluster and better understand the clustering process.

Step 2: Clustering Operator and Parameters

The modeling operator is available in the Modeling > Clustering and Segmentation folder, and is labeled DBSCAN. The allowing parameters can be configured in the model operator:

- Epsilon (ε): Size of the high-density neighborhood. The default value is 1.
- MinPoints: Minimum number of data objects within the epsilon neighborhood to qualify as a cluster.
- **Distance measure:** The proximity measure can be specified in this parameter. The default and most common measurement is Euclidean distance. Other options here are Manhattan distance, Jaccard coefficient, and cosine similarity for document data.

- **Add cluster as attributes:** To append cluster labels into the original data set. Turing on this option is recommended for later analysis.

Step 3: Evaluation (Optimal)

Similar to k-means clustering implementation, we can evaluate the effectiveness of clustering groups using average within cluster distance.

4. SYSTEM ARCHITECTURE

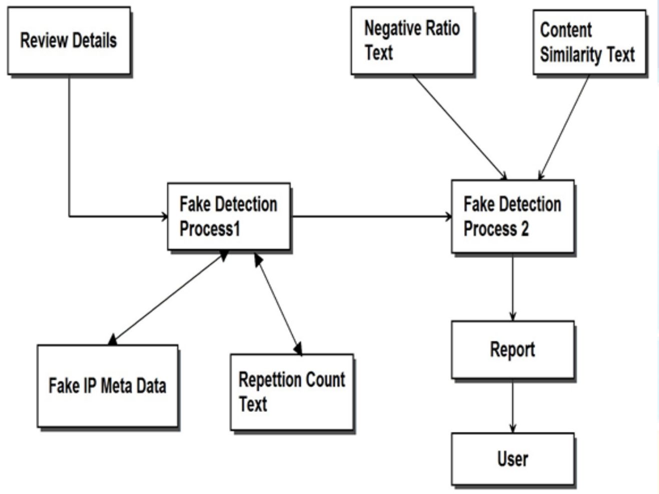


Figure 1 System Architecture

5. RESULTS

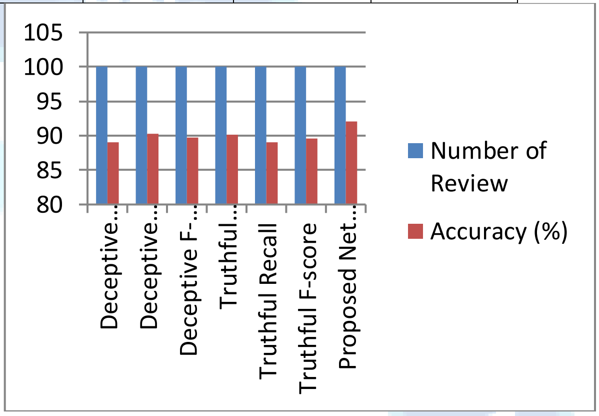
In this result chapter, we evaluate Spam detection from different perspective and compare it with two other approaches, Random approach and SPeaglePlus. To compare with the first one, we have developed a proposed system in which reviews are connected to each other randomly. Second approach use a well-known graph-based algorithm called as "LBP" to calculate final labels. Our observations show proposed system, outperforms these existing methods.

Then analysis on our observation is performed and finally we will examine our framework in unsupervised mode. Lastly, we investigate time complexity of the proposed framework and the impact of camouflage strategy on its performance.

1) Accuracy: Figures present the performance. As it's shown in all of the datasets proposed system outperforms SPeaglePlus specially when number of features increase. In addition different supervisions have no considerable effect on the metric values.

2) Table 1 Accuracy Comparison of Existing and Proposed System

SN	Name	Number of Review	Accuracy (%)
1	Deceptive Precision	100	89.1
2	Deceptive Recall	100	90.3
3	Deceptive F-score	100	89.7
4	Truthful Precision	100	90.1
5	Truthful Recall	100	89.0
6	Truthful F-score	100	89.6
7	Proposed Net Review Filter	100	94.6



neither on datasets proposed system nor SPeaglePlus. Results also show the datasets with higher percentage of spam reviews have better performance because when fraction of spam reviews in a certain dataset increases, probability for a review to be a spam review increases and as a result more spam reviews will be labeled as spam reviews and in the result of AP measure which is highly dependent on spam percentage in a dataset.

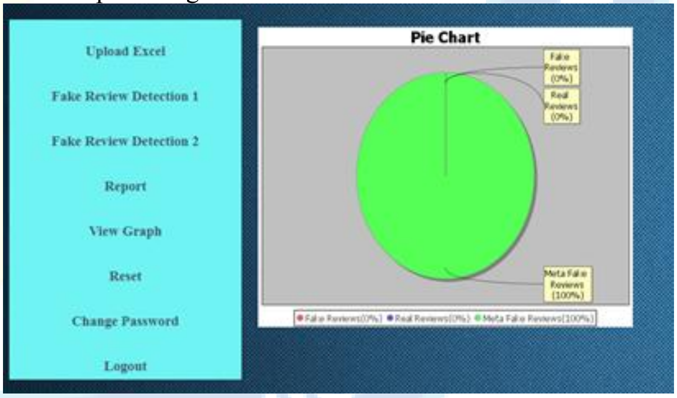


Figure 2. show detection of fake reviews in Pie Chart

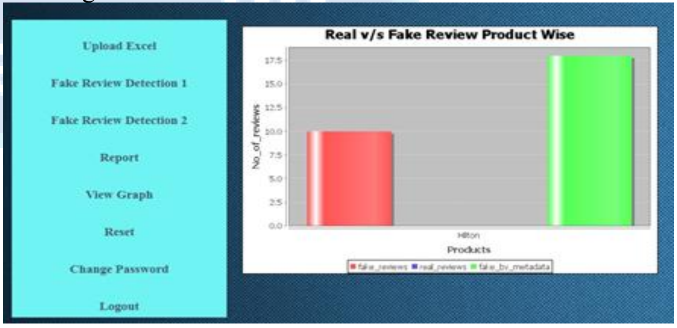


Figure 3. show the bar chart of Real v/s Fake Review product wise

6. CONCLUSION

NetSpam, a novel spam detection framework based on a meta path concept, and a novel graph-based method for labeling reviews using a rank-based labeling approach are both presented in this study. Two real-world labeled datasets from the Yelp and Amazon websites are used to measure the effectiveness of the proposed framework. Based on our observations, this meta path concept's calculated weights can be very effective at identifying spam reviews and improve performance. The meta path concept can be used to address issues in this field in future research. Similar frameworks, for instance, can be utilized to locate spammer communities. Reviews with the highest similarity based on the meta path concept are referred to as communities, and group spammer features (such as the proposed feature) can connect reviews to form communities. Additionally, as we utilized features more associated with spotting spammers and spam reviews, the utilization of product features represents an intriguing area of future research for this study. In addition, information diffusion and content sharing in multilayer networks is still in its infancy, despite the fact that researchers in a variety of fields have been paying close attention to single networks for more than a decade. One potential new area of study in this field is addressing the issue of spam detection in these networks.

REFERENCE

[1]. Heydari, M. A. Tavakoli, N. Salim, and Z. Heydari, (2014). Detection of review spam: A survey. Expert Systems with Applicants, Elsevier.

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

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

[3]. Mukherjee, B. Liu, and N. Glance, (2012), Spotting Fake Reviewer Groups in Consumer Reviews. In ACM WWW, 2012.

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

[5]. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah, K. P. Gummadi, B. Krishnamurthy, and A. Mislove, (2014). Towards detecting anomalous user behavior in online social networks. In USENIX.

[6]. L. Lai, K. Q. Xu, R. Lau, Y. Li, and L. Jing, (2011). Toward a Language Modeling Approach for Consumer Review Spam Detection. In Proceedings of the 7th international conference on e-Business Engineering.

[7]. Luo, R. Guan, Z. Wang, and C. Lin, (2014). HetPathMine: A Novel Transductive Classification Algorithm on Heterogeneous Information Networks. In ECIR.

[8]. Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features, (2014). In SIAM International Conference on Data Mining.

[9]. D. Wahyuni and A. Djunaidy, (2016). Fake Review Detection From a ProductReview Using Modified Method of Iterative Computation Framework. In Proceeding MATEC Web of Conferences.

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

[11]. Li, M. Huang, Y. Yang, and X. Zhu, (2011). Learning to identify review spam. Proceedings of the 22nd International Joint Conference on Artificial Intelligence; IJCAI.

[12]. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh, (2013). Exploiting burstiness in reviews for review spammer detection. In ICWSM.

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

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

[15]. H. Xue, F. Li, H. Seo, and R. Pluretti, (2015). Trust-Aware Review Spam Detection. IEEE Trustcom/ISPA .

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

[17]. Weise. A Lie Detector Test for Online Reviewers, (2016). http://bloom.bg/1KAxzhK. Accessed: 2016-12-16.

[18]. Akoglu, R. Chandy, and C. Faloutsos, (2013). Opinion fraud detection in online reviews bynetwork effects. In ICWSM.

[19]. Crawford, T. D. Khoshgoftar, J. N. Prusa, A. Al. Ritcher, and H. Najada, (2015). Survey of Review Spam Detection Using Machine Learning Techniques. Journal of Big Data.

[20]. Crawford, T. M. Khoshgoftaar, and J. D. Prusa, (2016). Reducing Feature set Explosion to Faciliate Real-World Review Sapm Detection. In Proceeding of 29th International Florida Artificial Intelligence Research Society Conference.

[21]. M. Luca and G. Zervas, (2016). Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud., SSRN Electronic Journal.

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

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

[24]. M. Salehi, R. Sharma, M. Marzolla, M. Magnani, P. Siyari, and D. Montesi, (2015). Spreading processes in multilayer networks. In IEEE Transactions on Network Science and Engineering. 2(2):65–83.

[25]. Jindal and B. Liu. Opinion Spam and Analysis, (2008). In WSDM.

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

[27]. R. Hassanzadeh, (2014). Anomaly Detection in Online Social Networks: Using Datamining Techniques and Fuzzy Logic. Queensland University of Technology, Nov.

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

[29]. S. Feng, L. Xing, A. Gogar, and Y. Choi, (2012). Distributional footprints of deceptive product reviews. In ICWSM.

[30]. Anurag et. al., “Load Forecasting by using ANFIS”, International Journal of Research and Development in Applied Science and Engineering, Volume 20, Issue 1, 2020

[31]. Raghawend, Anurag, "Detect Skin Defects by Modern Image Segmentation Approach, Volume 20, Issue 1, 2020

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

[33]. S. Mukherjee, S. Dutta, and G. Weikum, (2016). Credible Review Detection with Limited Information using Consistency Features, In book: Machine Learning and Knowledge Discovery in Databases.

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

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

[36]. Y. Sun and J. Han, (2009). Rankclus: integrating clustering with ranking for heterogeneous information network analysis. In Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology.

[37]. Y. Sun, J. Han, X. Yan, P. S. Yu, and T. Wu, (2011). Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. In VLDB.

.