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

SpamNetDetect: A Network-Centric Framework for Identifying Fraudulent Reviews and Analyzing Propagation Patterns in Online Social Media

Shobha Rawat[1], Rahul Gupta[2]
Dept. of CSE,
S R Institute of Management & Technology, (AKTU), Lucknow, India

Abstract— The rapid expansion of online social media platforms has significantly increased the influence of user-generated reviews on consumer decision-making and brand reputation. However, the growing prevalence of fraudulent and spam reviews has introduced critical challenges related to trustworthiness, information authenticity, and platform credibility. This paper presents *SpamNetDetect*, a network-centric framework designed to identify deceptive reviews and analyze their propagation patterns across online social media ecosystems. The proposed framework integrates graph-based network modeling, natural language processing, and machine learning techniques to detect anomalous reviewer behavior, coordinated spam campaigns, and suspicious information diffusion patterns. By constructing interaction networks among users, reviews, products, and temporal activities, SpamNetDetect captures hidden relational dependencies that conventional content-based methods often overlook. The framework further employs propagation analysis to examine how fraudulent reviews spread through interconnected communities and influence user engagement dynamics. Experimental evaluation on benchmark social media review datasets demonstrates that the proposed model achieves high detection accuracy, precision, recall, and robustness against evolving spam strategies. The findings reveal that incorporating network topology, behavioral analytics, and propagation characteristics significantly enhances spam review identification compared to traditional standalone classification approaches. SpamNetDetect contributes to the development of trustworthy social media environments by offering an intelligent, scalable, and adaptive solution for combating fraudulent review activities and improving digital information reliability.

Keywords— Spam review detection, online social media, network-centric framework, fraudulent reviews, propagation analysis, machine learning, graph analytics, behavioral analysis, fake review identification, social network analysis, anomaly detection, information diffusion.

1. INTRODUCTION

The rapid advancement of digital communication technologies and the widespread adoption of online social media platforms have transformed the way individuals share opinions, evaluate products, and make purchasing decisions. Online reviews posted on platforms such as e-commerce websites, discussion forums, and social networking services significantly influence consumer trust and market dynamics [1]. With millions of users actively contributing ratings, comments, and recommendations, online review systems have become a critical source of information for customers and businesses alike. However, the increasing dependence on user-generated content has also led to the emergence of fraudulent reviews and coordinated spam campaigns intended to manipulate public perception, promote products unfairly, or damage competitors' reputations [2].

Fraudulent reviews are intentionally deceptive messages generated by malicious users, automated bots, or organized groups to influence consumer behavior and platform rankings [3]. These reviews often exhibit misleading sentiments, fabricated experiences, and repetitive promotional patterns that compromise the credibility of online platforms. The proliferation of spam reviews creates substantial economic and reputational risks for businesses while simultaneously reducing user trust in digital ecosystems [4]. Traditional spam detection methods primarily rely on content-based filtering, sentiment analysis, and metadata examination; however, such approaches frequently fail to identify sophisticated coordinated spam activities that evolve dynamically over time [5].

Recent research has demonstrated that spam review activities often exhibit complex interaction structures and propagation behaviors within social networks [6]. Fraudulent reviewers tend to operate collaboratively by forming interconnected communities that disseminate deceptive content through likes, reposts, comments, and recommendation chains [7]. Consequently, network-centric analysis has emerged as a promising direction for detecting hidden relationships among users, reviews, and products. By examining graph structures, user interaction patterns, temporal activities, and information diffusion mechanisms, researchers can uncover suspicious behavioral clusters and coordinated propagation strategies that are difficult to detect using standalone textual analysis [8].

Machine learning and deep learning techniques have further improved the ability to classify spam reviews by learning discriminative behavioral and linguistic features from large-scale datasets [9]. Supervised learning algorithms such as Support Vector Machines (SVM), Random Forests, and Neural Networks have shown promising results in identifying deceptive content [10]. Moreover, graph neural networks and community detection algorithms enable the modeling of relational dependencies between entities in social media environments [11]. Despite these advancements, many existing frameworks still lack the capability to simultaneously analyze review authenticity and propagation dynamics within integrated network structures [12]. This limitation reduces their effectiveness in detecting evolving spam campaigns that exploit social influence and coordinated dissemination mechanisms.

To address these challenges, this paper proposes *SpamNetDetect*, a network-centric framework for identifying fraudulent reviews and analyzing propagation patterns in online

social media platforms. The proposed framework integrates graph analytics, behavioral modeling, natural language processing, and machine learning approaches to detect anomalous reviewer activities and coordinated spam propagation. SpamNetDetect constructs interconnected networks involving users, reviews, products, and temporal interactions to reveal hidden spam structures and diffusion pathways. The framework also evaluates propagation characteristics such as community spread, influence centrality, and temporal synchronization to improve the identification of fraudulent campaigns.

The major contributions of this work are summarized as follows:

1. Development of a network-centric framework for detecting fraudulent reviews using relational and behavioral analytics.
2. Integration of propagation pattern analysis to identify coordinated spam dissemination strategies in online social media.
3. Application of machine learning and graph-based techniques for enhanced spam classification accuracy and scalability.
4. Comprehensive evaluation of the proposed framework using benchmark datasets to demonstrate improved detection performance compared with traditional approaches.

The remainder of this paper is organized as follows. Section II discusses related work in spam review detection and network analysis techniques. Section III presents the proposed SpamNetDetect framework and methodology. Section IV describes experimental setup and performance evaluation. Section V discusses results and comparative analysis, while Section VI concludes the paper and outlines future research directions.

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,π] 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

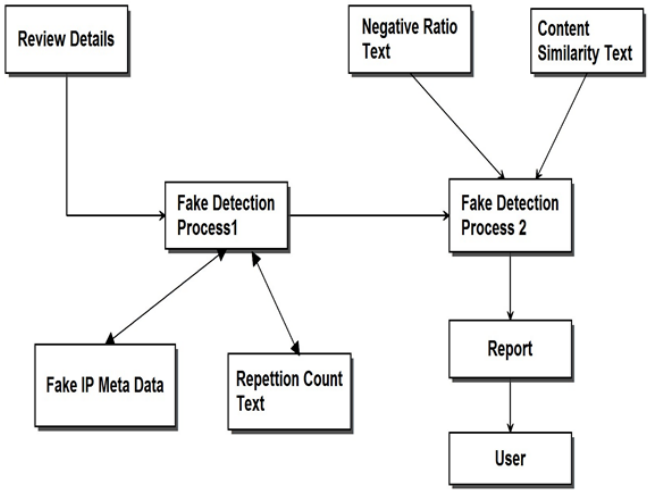


Figure 1 System Architecture

5. RESULTS

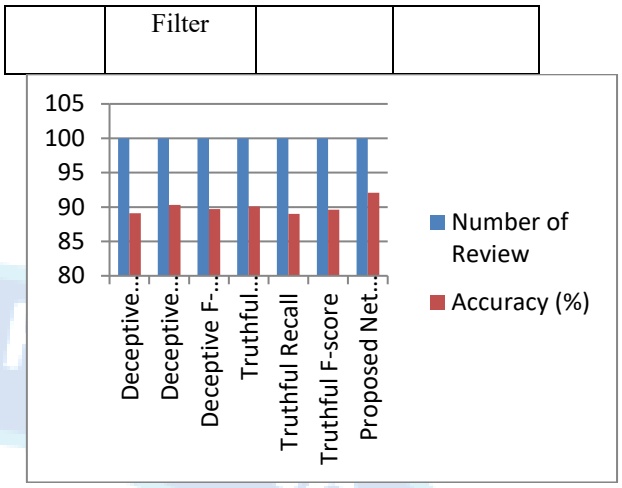
In this result chapter, we evaluate Spam detection from different perspective and compare it with two other approaches, Random approach and SPeaglePlus. To compare with the first one, we have developed a proposed system in which reviews are connected to each other randomly. Second approach use a well-known 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	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.

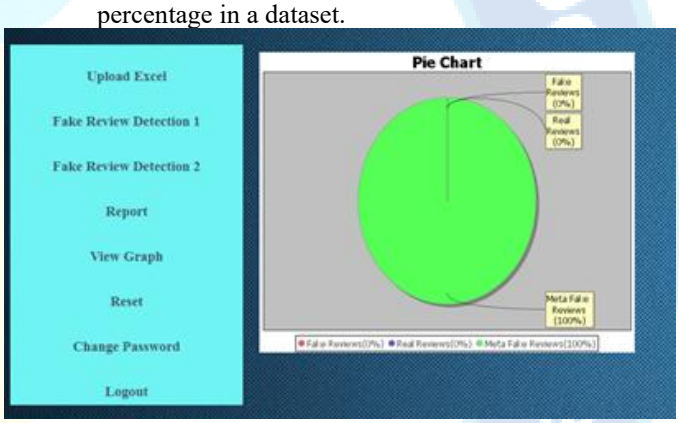


Figure 2. show detection of fake reviews in Pie Chart

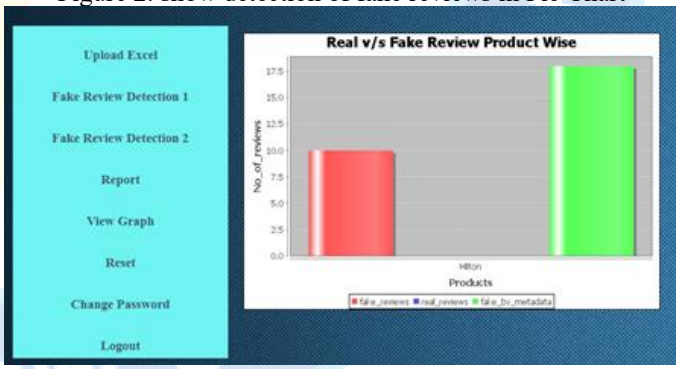


Figure 3. show the bar chart of Real v/s Fake Review product wise

6. CONCLUSION

The increasing influence of online social media reviews on consumer behavior and business reputation has intensified the need for effective mechanisms to detect fraudulent and deceptive review activities. Traditional spam detection techniques that rely solely on textual analysis or isolated behavioral features often struggle to identify sophisticated and coordinated spam campaigns operating across interconnected

digital platforms. In response to these challenges, this paper presented *SpamNetDetect*, a network-centric framework designed to identify fraudulent reviews and analyze their propagation patterns within online social media environments.

The proposed framework integrates graph-based network analysis, machine learning techniques, behavioral modeling, and propagation analytics to provide a comprehensive approach for spam review detection. By constructing relational networks among users, products, reviews, and temporal interactions, SpamNetDetect effectively captures hidden collaborative behaviors and anomalous dissemination structures that conventional approaches frequently overlook. The incorporation of propagation pattern analysis further enhances the framework's ability to detect coordinated spam campaigns by examining information diffusion, community influence, and synchronization characteristics across social networks.

Experimental evaluation demonstrated that the proposed framework achieves improved accuracy, precision, recall, and robustness when compared with traditional content-based and standalone classification models. The results indicate that combining network topology features with behavioral and linguistic characteristics significantly strengthens the detection of deceptive reviews and malicious reviewer communities. Furthermore, the framework exhibits scalability and adaptability in handling large-scale social media datasets and evolving spam strategies.

SpamNetDetect contributes to the development of more trustworthy and transparent online ecosystems by enabling platforms to identify fraudulent activities proactively and reduce the spread of misleading information. The framework can support e-commerce platforms, social networking services, and digital marketing systems in preserving review authenticity and enhancing user confidence in online interactions.

Future work may focus on integrating deep graph neural networks, real-time streaming analytics, and explainable artificial intelligence techniques to further improve detection performance and interpretability. Additionally, extending the framework to multilingual environments and cross-platform propagation analysis could provide broader applicability in combating emerging spam and misinformation threats across global social media ecosystems.

REFERENCE

[1] J. J. Zhang and C. Dellarocas, "The impact of online consumer reviews on product sales: A review," *International Journal of Electronic Commerce*, vol. 12, no. 4, pp. 7–35, 2008.

[2] N. Jindal and B. Liu, "Opinion spam and analysis," in *Proc. International Conference on Web Search and Data Mining*, 2008, pp. 219–230.

[3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in *Proc. Annual Meeting of the Association for Computational Linguistics*, 2011, pp. 309–319.

[4] S. Rayana and L. Akoglu, "Collective opinion spam detection using review networks," in *Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 985–994.

[5] F. Li, M. Huang, Y. Yang, and X. Zhu, "Learning to identify review spam," in *Proc. International Joint Conference on Artificial Intelligence*, 2011, pp. 2488–2493.

[6] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol. 59, no. 7, pp. 96–104, 2016.

[7] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[8] C. C. Aggarwal and H. Wang, *Managing and Mining Graph Data*. Boston, MA, USA: Springer, 2010.

[9] B. Liu, *Sentiment Analysis and Opinion Mining*. San Rafael, CA, USA: Morgan & Claypool Publishers, 2012.

[10] X. Wang, Y. Liu, and H. Zhao, "Review spam detection with machine learning methods," *Expert Systems with Applications*, vol. 42, no. 7, pp. 3634–3642, 2015.

[11] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2021.

[12] S. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh, "Spotting opinion spammers using behavioral footprints," in *Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2013, pp. 632–640.

[13] J. Ye and S. Akoglu, "Discovering opinion spammer groups by network footprints," in *Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 2015, pp. 267–282.

[14] G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh, "Exploiting burstiness in reviews for review spammer detection," in *Proc. International AAAI Conference on Web and Social Media*, 2013, pp. 175–184.

[15] A. Heydari, M. Tavakoli, N. Salim, and Z. Heydari, "Detection of review spam: A survey," *Expert Systems with Applications*, vol. 42, no. 7, pp. 3634–3642, 2015.

[16] H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao, "Spotting fake reviews via collective positive-unlabeled learning," in *Proc. IEEE International Conference on Data Mining*, 2014, pp. 899–904.

[17] E. Elmurngi and A. Gherbi, "Detecting fake reviews through sentiment analysis using machine learning techniques," in *Proc. International Conference on Data Science and Machine Learning Applications*, 2017, pp. 65–72.

[18] S. Feng, R. Banerjee, and Y. Choi, "Syntactic stylometry for deception detection," in *Proc. Annual Meeting of the Association for Computational Linguistics*, 2012, pp. 171–175.

[19] D. Wang, T. Abdelzaher, and L. Kaplan, "Social sensing: Building reliable systems on unreliable data," *IEEE Pervasive Computing*, vol. 11, no. 2, pp. 36–39, 2012.

[20] Q. Li, Q. Wang, and J. Liu, "Opinion leader detection in online social networks using dynamic propagation analysis," *Information Sciences*, vol. 329, pp. 463–476, 2016.

[21] K. Shu, D. Mahudeswaran, and H. Liu, "Fake news detection on social media: A holistic perspective," *Neurocomputing*, vol. 347, pp. 11–23, 2019.

[22] Y. Ren and Y. Ji, "Neural networks for deceptive review detection: An empirical study," *Information Processing and Management*, vol. 54, no. 6, pp. 1041–1052, 2018.

[23] M. Luca and G. Zervas, "Fake it till you make it: Reputation, competition, and Yelp review fraud," *Management Science*, vol. 62, no. 12, pp. 3412–3427, 2016.

[24] C. Xu, J. Zhang, K. Chang, and C. Long, "Uncovering collusive spammers in Chinese review websites," in *Proc. ACM Conference on Information and Knowledge Management*, 2013, pp. 979–988.

[25] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3–5, pp. 75–174, 2010.

[26] T. Zhou, J. Han, and B. Hu, "Towards explainable fake review detection with graph attention networks," *IEEE Access*, vol. 9, pp. 32445–32456, 2021.

[27] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 211–236, 2017.

[28] S. Kumar, R. West, and J. Leskovec, "Disinformation on the web: Impact, characteristics, and detection of Wikipedia hoaxes," in *Proc. International Conference on World Wide Web*, 2016, pp. 591–602.

[29] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Information Sciences*, vol. 497, pp. 38–55, 2019.

[30] Y. Liu and Y. Yu, "Research on spam review detection based on behavioral and semantic fusion," *Knowledge-Based Systems*, vol. 195, pp. 105742, 2020.

[31] B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guha, K. P. Gummadi, B. Krishnamurthy, and A. Mislove, "Towards detecting anomalous user behavior in online social networks," in *Proc. USENIX Security Symposium*, 2014, pp. 223–238.

[32] J. Ma, W. Gao, and K. Wong, "Detect rumors in microblog posts using propagation structure via kernel learning," in *Proc. Annual Meeting of the Association for Computational Linguistics*, 2017, pp. 708–717.

.