SPAM DETECTION USING SEMANTIC AND TEMPORAL ANALYSIS IN REVIEWS

A DISSERTATION

submitted in partial fulfillment of the requirements

for the award of the degree of

MASTER OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

By

Khushbu Jhunjhunuwala

17535009

Under the supervision of

Dr. Durga Toshniwal

Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY, ROORKEE

ROORKEE - 247667

May, 2019

CANDIDATE'S DECLARATION

I hereby declare that the dissertation entitled **"Spam Detection using Semantic and Temporal Analysis in Reviews"** submitted by me in partial fulfilment of the requirements for the award of the Degree of Master of Technology in Computer Science and Engineering to the Department of Computer Science and Engineering, Indian Institute of Technology Roorkee is my original work carried during August 2018 to April 2019 under the guidance of **Dr. Durga Toshniwal, Professor**, Department of Computer Science and Engineering, Indian Institute of Technology, Roorkee.

The content presented in this dissertation has not been submitted by me for award of any other degree of this and any other institute.

Date:

Place: Roorkee **Khushbu Jhunjhunuwala**

CERTIFICATE

This is to certify that Thesis Report entitled **"Spam Detection using Semantic and Temporal Analysis in Reviews"** which is submitted by **Khushbu Jhunjhunuwala (17535009)** towards the fulfilment of the requirements for the award of the Degree of **Master of Technology in Computer Science & Engineering**, submitted to the Department of Computer Science & Engineering, **Indian Institute of Technology Roorkee**, India is carried out by her under my esteemed supervision and the statement made by the candidate in the declaration is correct to the best of my knowledge and belief.

Date:

Place: Roorkee

Dr. Durga Toshniwal

Professor

Dept of Computer Science & Engineering

Indian Institute of Technology, Roorkee

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude towards supervisor Dr. Durga Toshniwal for the continuous support, motivation, and guidance. I would also like to thank my lab-mates and friends for their support and help.

I am also grateful to the Department of Computer Science & Engineering of IIT Roorkee for providing valuable resources to aid my research.

Khushbu Jhunjhunuwala

ABSTRACT

The opinions on online platforms like Amazon, Goibibo, TripAdvisor for products or services are widely used by customers or users for their decision making in recent years. The products or services which are highest rated attract maximum attention of users and are most likely to get purchased. Looking this trend on e-commerce sites, spammers deceive users intentionally by giving dishonest reviews of products to give undue promotion for their products and demote the products of their competitors. The existing state-of-the-art techniques has done behavioral analysis on the features, graphical analysis on review or reviewer or product relationships, other supervised learning approaches to identify spam reviews. There is still a lot scope to work on the temporal and semantically similar behavior among reviews.

This thesis work has been taken to explore the temporal behavior and the semantic similarity of reviews and identify the unusual high deviation patterns. Some active zones which spammers adopt are identified which further depend on average truthful ratings of the product. Similarity analysis reveals the existence of a similarity range which spam reviews show and can be used to identify reviews as spam or genuine. Many ways to capture the similarity are tried and checked if this can help reduce false positives. A hybridization of both these analysis again proves the existence of this behavior of spammers.

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

1.1 Introduction and Motivation

User-generated content is becoming increasingly valuable to both individuals and businesses due to its importance and influence in e-commerce markets. Both customers and businesses have embraced online platforms for their conveniences, better deals, more varieties and other advantages. With more than 3 billion people using Internet today, they post their reviews for the products or services they use on these platforms or discuss problems on the online forums. These reviews have immense power to impact people and shape their decisions. Customers read these reviews before purchasing the product and take their decisions based on them. If the review rating is higher and most of the reviews posted from other buyers are positive, the probability of customer purchasing the product increases. Similarly, if the rating for the product is not good and the reviews posted are negative, the willingness of customer buying that product decreases and he decides to purchase some other product. Thus, reviews play a pivotal role for the promotion or demotion of businesses. However, the genuineness of these reviews is nowhere mentioned. This leads to opinion spamming. Opinion spamming is writing fake reviews for products or services to mislead users intentionally by providing fake reviews to some products to boost their sell or to destroy their name. This fraud is employed by many businesses for their name and fame. Opinion spamming can be in the form of fake

reviews, dishonest up votes or feedbacks on reviews, likes on Facebook pages, fake YouTube subscribers, Twitter followers etc. No verification and anonymous nature of reviewers help these spammers to influence the reviews.

Opinion Spamming is on the hike every day. The below two examples from the news prove how this has affected the online users.

- A Daily Mail report [1] shows this problem on TripAdvisor as how reviews on this website are spammed by users. Reviews are being sold for as cheap as $3. These spammers are normal people who deliberately write dishonest reviews for businesses for very small charges. Many businesses, other service owners let spammers work for them by writing fake positive reviews for their services on Trip Advisor and defame their competitors business. This proves that these online platforms or social networking websites do not perform any verification for reviews when posted online to catch fraudsters. These reviews on ecommerce websites influence the minds of actual customers purchasing the products. Online websites have become an important part of trading policies to attract customers.
- One more scene reported by CNBC [2] tells how these online platforms are liable to online cheat. "Pure Daily Care" is a skin care products selling company. It became the victim of opinion spam and the sales of its large selling products degraded in a month by the flood of negative feedbacks on Amazon. The spammers thus act dishonestly by influencing the average ratings and defaming the real true products. This opinion spam cause the financial loss to the potential customers as well as to the businessmen. Also it destroys the trust of the users on the online platforms.

Figure 1.1 shows an example where an exactly same review text has been written for two different products on same date.

The below two examples shows similar reviews written by same reviewer.
Example 1:

- *We absolutely love this place.Service is always outstanding with experienced servers.*

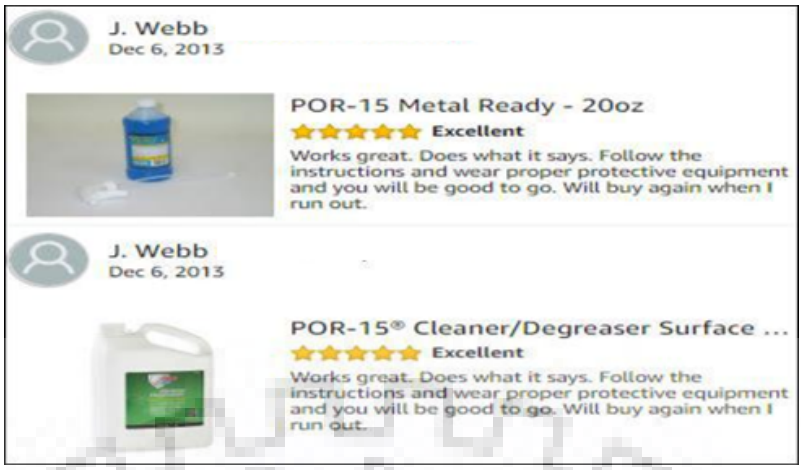

Figure 1.1: Identical texts for two different products [14] [15]

- *Hands down my favorite restaurant. Angelo and Enzo are great and always so welcoming-like you are coming to their home.*

Example 2:

- *BEST italian joint on the Northside. The food is GREAT.*
- *If you are a true Italian food fan this sits up there with the best.*

Similarly, temporal analysis show bursts among reviews which shows the possibility of spam.

These online social platforms must be well equipped with the anti-spam detection techniques to get rid of any type of fraud present in the reviews to maintain the interests of people and show interactions in their real form.

1.2 Problem Description

The spammers are adopting new and smart strategies to post genuine looking reviews. Spammers could also write some sincere reviews to get into the group of truthful reviewers and mask themselves. The problem focuses on detecting these fake reviews on online platforms by analyzing the deviations of review characteristics from the general behavior.

The proper analysis on the characteristics or features of reviews has to be done to identify various patterns that exhibit in the reviews written by spammers and identify the strategies adopted by spammers. Reviews are then to be classified as spam or non-spam.

Given a dataset containing review tuples R, the problem is to identify opinion fraud reviews i.e a prediction function f that labels the reviews:

$$\text{f} : \text{R} \rightarrow \{fraud, genuine\}.$$

Assumptions: Spammers do not show 100% similarity. This much high similarity is trivial to identify and exist hardly. This work focuses on identifying spam based on semantic similarity.

1.3 Specific Research Contributions

This work is based on analyzing the temporal and semantic characteristics which is found to exist in spam reviews. Temporal analysis includes analyzing the burst intervals which identifies the existence of different active zones of spammers based on the actual cumulative ratings. Semantic analysis is done on the review texts and is based on identifying the similarity among spam reviews. Multiple similarity algorithms are identified which can capture the semantic essence among reviews. Spammers are found to show relatively high similarity than the genuine reviewers. Both these analysis, when brought together, proves the existence of this behavior of spammers.

1.4 Organization of Report

The report is organized into five sections. The first section gives a brief introduction about opinion spam filtering and discusses the motivation to choose this topic. The second section describes the related work which has been done in the field of opinion spam filtering and also discusses research gaps. The third section describes the proposed methodology. The fourth section describes the experiments performed and the analysis done on the results. The fifth section concludes the report and discusses the future work.

Chapter 2

RELATED WORK

2.1 Literature Survey

The first work in this area was proposed by Jindal and Liu [3]. They analyzed patterns based on the number of reviews written by reviewers, number of reviews written for a product and identified an existing power law distribution among that. Some products are reviewed highly and some products have a very less number of reviews associated with them. Similar is the case with reviewers. The review texts were analyzed to identify the duplication among them. Jaccard distance was used a similarity measure. Duplicates were used as an important feature for model building. This work was remarkable but had shortcomings like targeting only a subset of opinion spam i.e. duplicate reviews. Moreover, spammers adopt very smart strategies to adulterate reviews which are difficult to identify using only this idea. In [4], Mukherjee et al. worked on identifying groups of spammers who work collaboratively on a group of products or a single product. Group spammers' were identified based on their behavioral patterns like group time window, group deviation, group content similarity, group time frame and other group oriented features. Li et al. in [5] identified the existence of bimodal distribution in the posting of reviews. There exist a specific time when these group of spammers become active and start posting online while the original reviewers do not exhibit any specific time of posting but any random time. The above work was based on the supervised learning approaches using

behavioral and linguistic patterns of reviewers. The work in [6] incorporated the semantic similarity of reviews. Frame rate and Bi-frame rate were proposed based on the statistical analysis of the semantic features of frames. It was found that true reviews contain the exact behaviors or their real encounter with the products or services like the size, type of the object in their reviews. In [7], Sandulescu et al. used semantic similarity to identify singleton spammers. Singleton spammers are those who write one review and then disappears. Spammers tend to write only a single review under a name for a product to avoid being caught. However, assuming that a person utilize a certain set of words repeatedly, semantic similarity was considered as a measure to identify them. Heydari et al. in [8] identified the existence of temporal patterns in the reviews. They constructed time series based on the number of reviews in the given time interval. The time period containing extraordinary number of reviews were identified as suspicious time intervals. Other behavioral features were then analyzed of these reviews to decide a spam score for these reviews. Mukherjee et al. in [9] again worked on temporal dynamics using the Yelp Dataset. Most of the works done employed manual filtering of the reviews relying on two or more than two experts in this field. This work used Yelp's filtered and recommended reviews. They identified that there was buffered spamming for entities that required spamming to retain their popularity and reduced spamming for others who were rated well by users. Using this analysis, they leveraged their work with the idea of pre detection of deception. They designed an auto vector regression model on the time series of the number of reviews to predict the spamming policy which is going to be adopted for a specific product/service. In [10], Siddu et al. constructed multi-dimensional time series based on the number of reviews, average ratings, positive word length score and negative word length score. If there is a burst detected in more than two dimension of time series were considered as suspicious time intervals. Also if any one of the dimension shows different patterns than other three were identified as abnormal patterns. In [13], they studied the performance of neural networks on the spam dataset and identified that they perform better on the cross-domain dataset. Gated Recurrent Neural Network are found to perform better with spam review dataset.

2.2 Research Gaps

Most of the work done so far is based on the behavioral features of reviews and the reviewers. These features like rating deviation, number of reviews, review sentiment, are time variant. Identifying the bursts and then analyzing the spamming policies have much scope for work.

Opinion spam is found in the reviews of products as well as services. The reviews for both of them can behave very differently. Some products like electronic items have a lifespan of few years but the services like restaurants, hotels have a larger lifespan. So the spammers can behave more smartly with the reviews of these services and are difficult to identify.

The identical reviews written by spammers using different ids can be identified by doing a semantic analysis on the review dataset. Every human has a set of vocabulary which they use or they write synonyms of the words they use. If two reviews are written by the same spammer, the similarity between those reviews is likely to be more than if they are written by two different spammers. This can be used to identify the spamming nature of reviewers.

Chapter 3

PROPOSED WORK

This work is based on analyzing the temporal behavior and semantic similarity existing in the review dataset. The preprocessing of dataset is done which includes stopwords removal, lemmatization of texts, product-wise segregation of review dataset and the analysis is carried out.

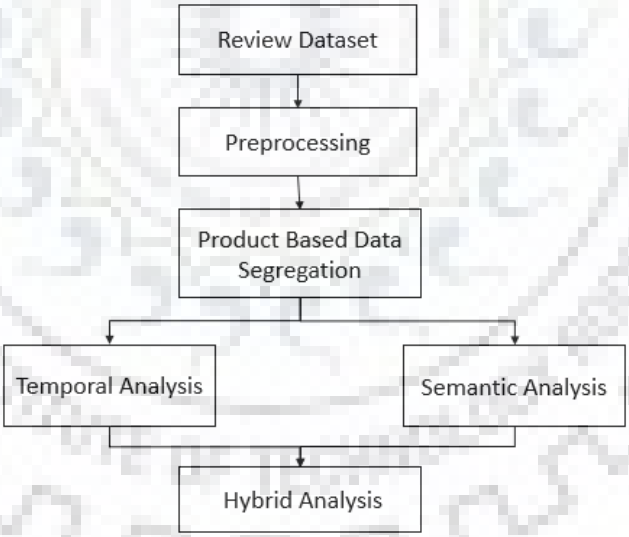

Figure 3.1: Methodology

3.1 Temporal Analysis

The temporal analysis includes creating a multi-dimensional time series and identifying the major active zones of spammers. A further segregation of deceptive and

non-deceptive reviews is done and the deceptive reviews are analyzed by constructing a multi-dimensional time series over its features. This is done to identify the specific strategies that spammers adopt to spam the reviews. The time series are then analyzed and some major active zones of spammers are identified which are based on the previous average cumulative ratings of the products. Their spamming policies vary as per the hotel review counts and the ratings. A K-means clustering algorithm is implemented to cluster the hotels.

3.2 Semantic Analysis

This analysis is based on identifying the similarity among reviews. Based on this analysis, it will construct a Spam Detection System to classify review as spam or non-spam.

3.2.1 Identifying Similarity among Reviews

The Bag-of-Words, tf-idf model are generally used to convert a document into sentence vectors. However these models do not consider the semantic meaning of words in the sentences. Six different algorithms to identify the similarity between reviews are identified which are named as Doc2Vec model and cosine similarity with lemmatized tokens(D2VL) ,Doc2Vec model and cosine similarity with non-lemmatized tokens(D2VNL), Word2Vec and cosine similarity with lemmatized tokens(W2VL), Word2Vec and cosine similarity with non-lemmatized tokens(W2VNL), MihalCea et al. algorithm [11] with lemmatized tokens(MCL), MihalCea et al. algorithm [11] with non-lemmatized tokens(MCNL). These algorithms will be denoted by the abbreviations above henceforth.

The segregated reviews are extracted and tokenized. The stop-words are removed and the remaining tokens are lemmatized. The similarity among each hotel review texts is identified using these algorithms. The Doc2Vec and Word2Vec models are run on the large review text to get the sentence vectors. From the large text corpus, the reviews would be better able to identify the semantic meaning of words in them. The cosine similarity among the pairwise reviews of each hotel is calculated. The MihalCea et al.[11] algorithm is also implemented to calculate pairwise similarity of

reviews of each hotel.

3.2.2 Label review as spam or genuine

For each hotel, reviews are labeled as spam or non-spam based on a similarity threshold. If the similarity is higher, reviews are labeled as spam. The recall and F1-Score for each threshold is analyzed. Based on this, a similarity range is identified for each algorithm which gives the maximum F1-Score.

3.3 Hybrid Analysis

A hybrid analysis using both the semantic and temporal of reviews is done. From the time series created in the temporal analysis, the reviews of the bursty intervals are extracted. The bursty intervals are taken as the intervals where the number of reviews are higher. The similarity among these reviews is extracted using different similarity algorithms to identify the similarity score exhibited and analyzed. It proves the existence of similarity score range among the reviews of spammers.

Chapter 4

EXPERIMENTS AND DISCUSSION

4.1 Dataset Used

The dataset used is Yelp Hotel Dataset. It contains reviews from 129 hotels from Chicago area. The recommended reviews of Yelp are considered as genuine reviews and the filtered ones are considered as spam reviews. Each tuple holds reviewer id, hotel id, review id,review timestamp, rating and review text.

Table 4.1: Dataset Details

Total number of Reviewers	33500
Total number of Hotels	129
Total number of Reviews	61538
Total number of Spam Reviews	8141
Total number of Genuine Reviews	53397

4.2 Evaluation Metrics

The dataset used is class unbalanced. So accuracy is not used as an evaluation metric of results. The Recall and F1-score are used as evaluation metrics.

Precision: It is defined as the proportion of identified spam which is actually correct as in Equation 4.1

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} = \frac{TruePositive}{TotalPredictedPositive} \tag{4.1}$$

Recall: It is defined as the proportion of actual spam identified correctly as in Equation 4.2.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} = \frac{TruePositive}{TotalActualPositive} \tag{4.2}$$

F1-score: It is given as the harmonic mean of both of them as in Equation 4.3.

$$F1Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4.3}$$

4.3 Results And Analysis

4.3.1 Temporal Analysis

It is assumed that spammers are more interested in increasing the spam ratings of hotels rather than decreasing the average ratings. A multi-dimensional time series is constructed for each hotel. K-means clustering algorithm is implemented to identify the number of hotels falling in each active zone of spammer. Table 4.2 lists the active zones of spammers identified based on the analysis.

Table 4.2: Identified Active Zones of Spammers

Active Zones	**Number of Hotels**
Late activeness of spammers	25
Mid activeness of spammers	35
Early activeness of spammers	17

Review count showing activeness of spammers at early stage is shown in Figure 4.1 for two hotels. The spike in the early stage can be seen in the time series. Review count showing activeness of spammers at late stage is shown in Figure 4.2 for two hotels. The spike in the late stage can be seen in the time series. Figure 4.3 shows

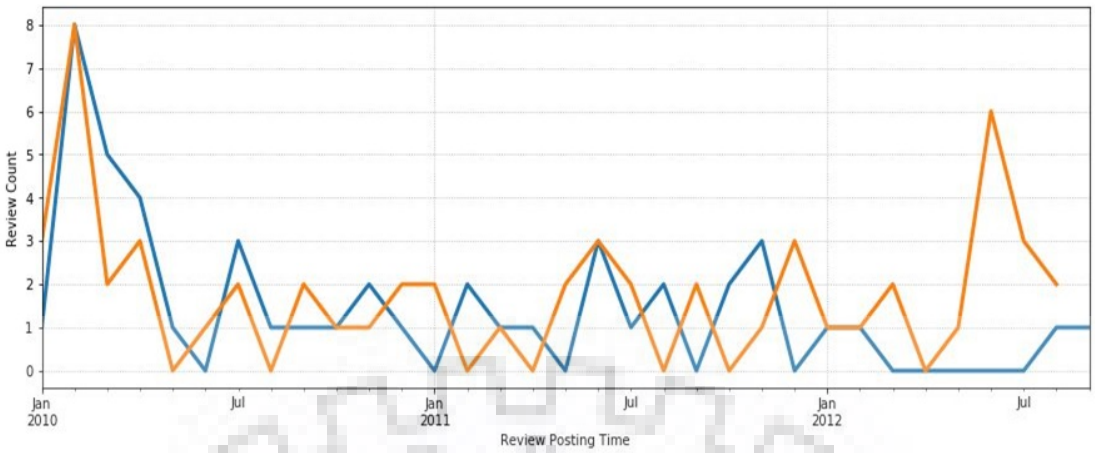

Figure 4.1: Review Count showing Activeness of Spammers at Early Stage for two hotels

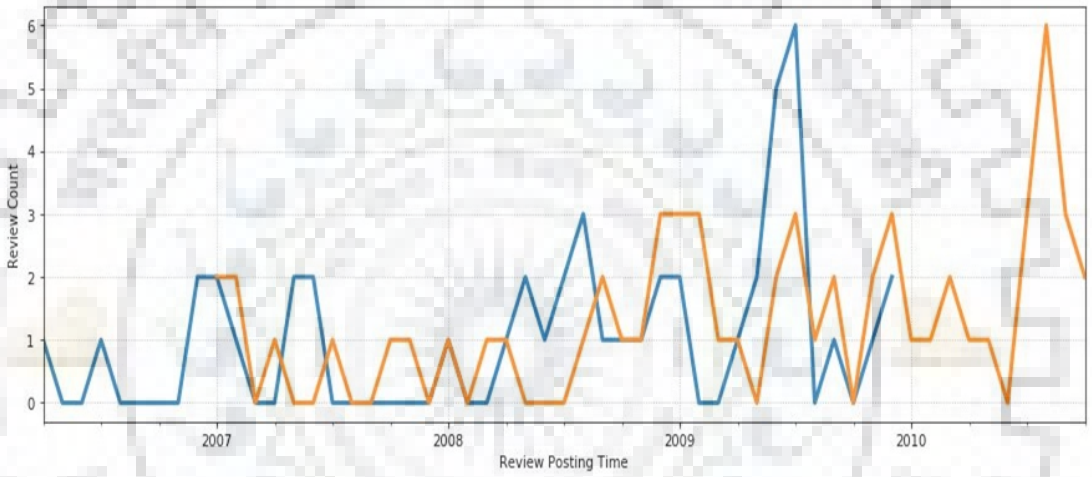

Figure 4.2: Review Count showing Activeness of Spammers at Late Stage for two hotels

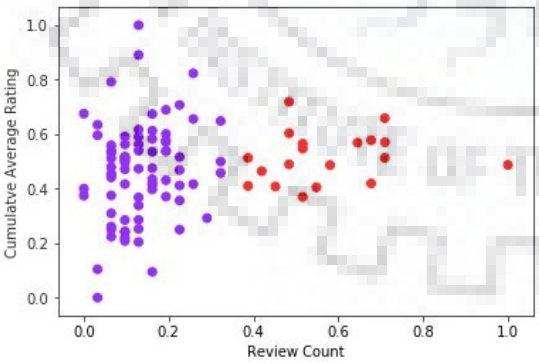

Figure 4.3: Clustering of hotels to identify early spammed hotels

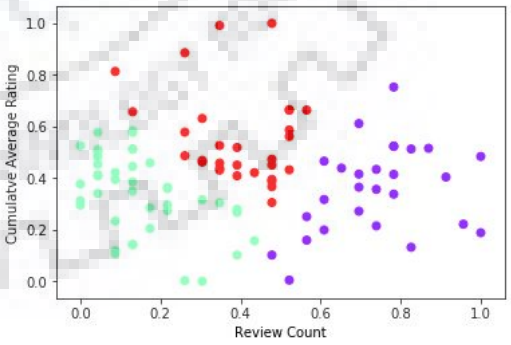

Figure 4.4: Clustering of hotels to identify late spammed hotels

the K-means clustering which is done to get the count of hotels which lie in the early spammed active zone. Figure 4.4 shows the K-means clustering which is done

to get the count of hotels which lie in the late spammed active zone. This behavior of spammers is apparent as the activeness of spammers depend on the real truthful ratings of hotel.

4.3.2 Semantic Analysis

The similarity is computed among the reviews as per the six different algorithms and the algorithm to label reviews as spam or genuine is run. The Doc2Vec and the Word2Vec model is trained with this large review text corpus for 100 epochs with a surrounding word length of 10.Three different vector size of 20, 50, 100 are tried. The threshold is varied from a range of 0.1 to 0.9.

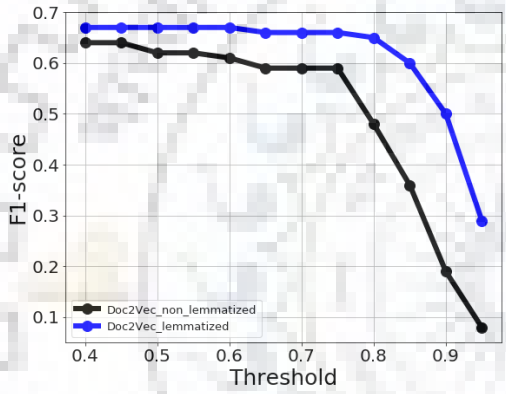

Figure 4.5: F-1 Score using Doc2Vec algorithm with cosine similarity

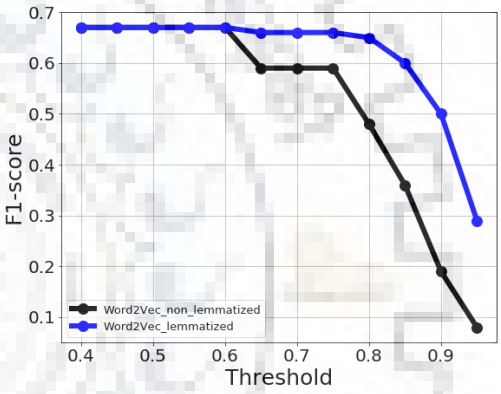

Figure 4.6: F-1 Score using Word2Vec algorithm with cosine similarity

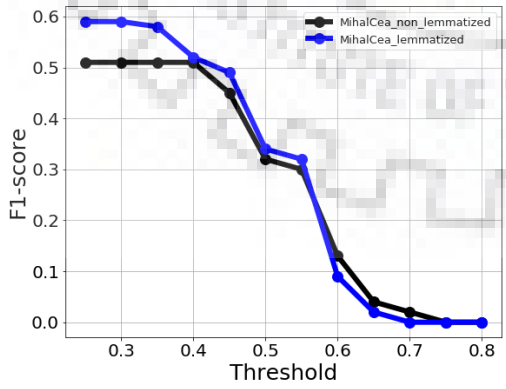

Figure 4.7: F-1 Score using MihalCea et al. [11] algorithm to identify similarity

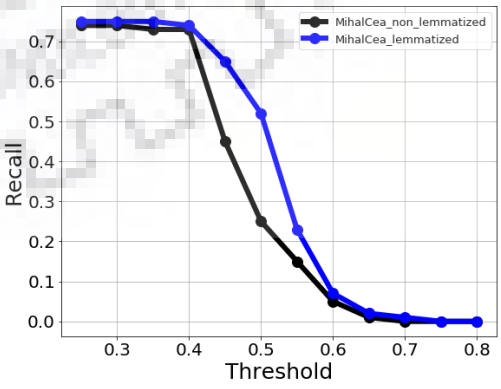

Figure 4.8: Recall using MihalCea et al. [11] algorithm to identify similarity

The F1-Score obtained on varying the threshold range using the Doc2Vec and Word2Vec algorithm for their two variations, the lemmatized tokens and the non-lemmatized tokens is shown in Figure 4.5 and 4.6. The F1-Score obtained on varying the threshold range using the MihalCea et al.[11] algorithm for their two variations, the lemmatized tokens and the non-lemmatized tokens is shown in Figure 4.7. Figure 4.8 shows the Recall obtained on varying the threshold range using the MihalCea et al.[11] algorithm for the two variations. The Recall obtained on varying the threshold range using the Doc2Vec and Word2Vec algorithm for their two variations, the lemmatized tokens and the non-lemmatized tokens is shown in Figure 4.9 and 4.10. The results from the above graphs can be summarized in Table 4.3.

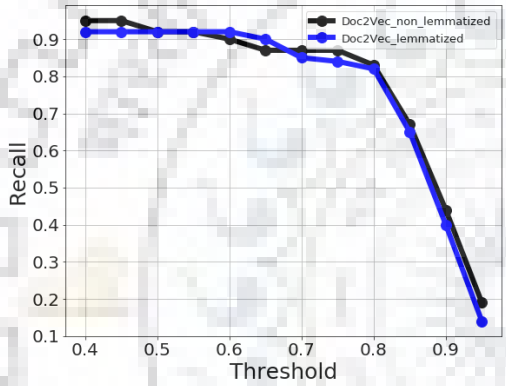

Figure 4.9: Recall using Doc2Vec algorithm with cosine similarity

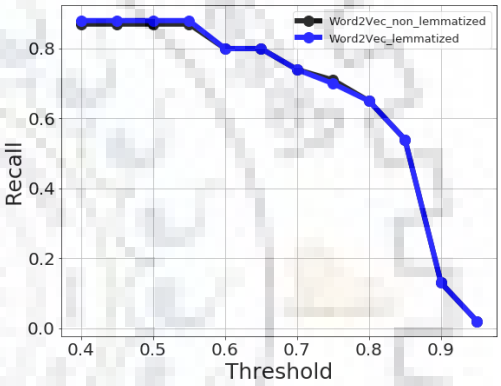

Figure 4.10: Recall using Word2Vec algorithm with cosine similarity

Table 4.3: F-1 Score, Recall, Spam Similarity Range for similarity algorithms

Similarity Algorithm	**F1-Score**	**Recall**	**Precision**	**Spam Similarity Range**
D2VL	**0.67**	**0.95**	0.53	0.45 - 0.8
D2VNL	0.64	0.92	0.49	0.45 - 0.6
W2VL	**0.67**	0.88	0.53	0.4 - 0.6
W2VNL	**0.67**	0.87	0.54	0.4 - 0.6
MCL	0.59	0.75	0.49	0.3 - 0.35
MCNL	0.51	0.74	0.38	0.3 - 0.4

Comparing the results in Figure 4.11 with the work done in [7], who also worked on

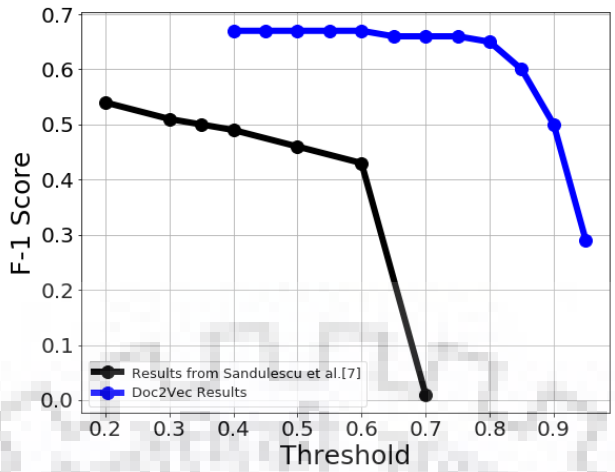

Figure 4.11: Comparison with Sandulescu et al. [7]

identifying the similarity among reviews on Yelp Dataset and using that to propose a Spam Detection System, the results outperforms these baselines. The highest F1-Score achieved is 0.67, which is higher compared to them of 0.55(approximately). Their recall is low(exact number was not specified) whereas this work shows a high recall of 0.95 at the threshold of 0.65 when similarity is found using D2VL.

To get better insight into how the similarity varies for the spam and genuine reviews,

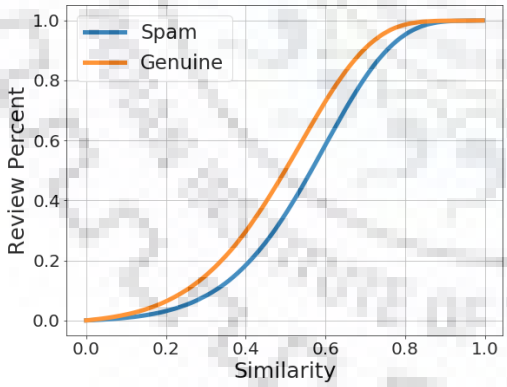

Figure 4.12: Review Count Distribution for spam/genuine reviews using Doc2Vec with cosine similarity algorithm

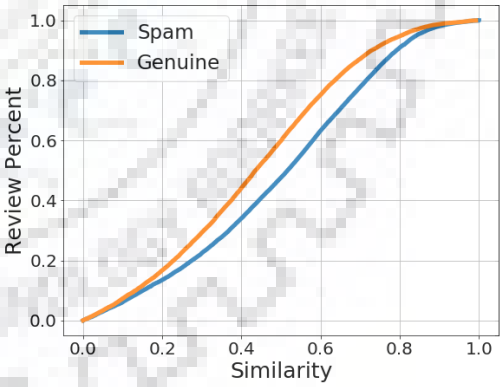

Figure 4.13: Review Count Distribution for spam/genuine reviews using MihalCea et al. [11] similarity algorithm

the similarity among the positive and negative reviews using the Doc2Vec with cosine similarity algorithm and the MihalCea et al. [11] algorithm is calculated and a cumulative distribution curve among the review count and their pairwise

similarity is plotted. The graphs in Figure 4.12 and 4.13 show that irrespective of the algorithm used to calculate spam threshold similarity among reviews, the spam reviews show a higher similarity than the genuine reviews. This proves as a great example to prove that similarity can be an important factor to identify the spam among reviews.

4.3.3 Hybrid Analysis

A hybrid analysis using the temporal features and semantic similarity of review texts is done. The semantic similarity of the reviews in bursty intervals of different spammer active zones using different similarity algorithms is identified. Table 4.4 shows the review pair count and their similarity score range using different similarity algorithms. It can be seen easily that the maximum review pair count lies in the similarity range identified above in all the three similarity algorithms. A large review pair count also exists in a higher similarity range as identified above. This can be because this similarity is found among the reviews identified in bursty intervals. The possibility of these reviews written by a spammer is comparatively higher than that of other reviews.

Table 4.4: Review Pair Count and their similarity range using different similarity algorithms

Similarity Range	**Review Pair Count**		
	Doc2Vec	**Word2Vec**	**MihalCea et al. [11]**
[0.0 - 0.1]	126	308	216
[0.1 - 0.2]	269	606	900
[0.2 - 0.3]	590	994	4,196
[0.3 - 0.4]	1,097	1,716	**11,674**
[0.4 - 0.5]	1,975	2,654	6,580
[0.5 - 0.6]	3,854	3,930	532
[0.6 - 0.7]	5,646	5,072	24
[0.7 - 0.8]	**5,913**	**5,394**	2
[0.8 - 0.9]	3,982	2,948	0

Chapter 5

CONCLUSION AND FUTURE WORK

Opinion Spamming is a nuisance to all the online social platforms which includes food, hotel, travel etc. platforms. It spoils the very aim of review systems of e-commerce sites to help their users deciding the quality of products. It is difficult to prevent opinion fraud and focus must be towards detecting and eradicating the malignant entities.

This work analyses the temporal characteristics and semantic similarity of reviews of spam and gives a insight into the strategies which spammers follow to ruin the truthful ratings of hotels. The temporal analysis reveal the existence of five active zones of spammers on hotels which depend on the actual rating behavior by the users. In semantic analysis, six different models to capture the semantic similarity among reviews are identified and it is shown that the similarity of spam reviews is higher than those of non-spam reviews irrespective of the model used. Doc2Vec and Word2Vec models give the best results and outperform the baselines set in the semantic analysis. A hybrid model using both these behaviors again proves the existence of this behavior of spammers.

The future work can be identifying better ways to capture the semantic similarities in reviews. Other features of reviews can be used along with semantic similarity captured to extract the feature vector of reviews. Research can be done using other techniques which does not require already annotated dataset.

REFERENCES

[1] T. Rawstorne, 'Disturbing proof the online review that made you book your holiday may be FAKE: Investigation reveals an entire industry is dedicated to generating bogus appraisals for cash', DailyMail, 2015.

[2] A. Levy, 'This Amazon seller lost $400,000 in sales after being attacked by self-proclaimed 'virus of Amazon', CNBC, 2017.

[3] N. Jindal, and B. Liu. Opinion spam and analysis. In the Proceedings of the 2008 International Conference on Web Search and Data Mining, WSDM '08, pages 219-230, 2008.

[4] A. Mukherjee, B. Liu, and N. Glance. Spotting fake reviewer groups in consumer reviews. In the Proceedings of the 21st international conference on World Wide Web, WWW '12, pages 191-200, 2012.

[5] H. Li, G. Fei, S. Wang, B. Liu, W. Shao, A. Mukherjee, and J. Shao. Bimodal Distribution and Co-Bursting in Review Spam Detection. In the Proceedings of the 26th International Conference on World Wide Web, WWW'17, pages 1063-1072, 2017.

[6] S. Kim, H. Chang, S. Lee, M. Yu, and J. Kang. Deep Semantic Frame-Based Deceptive Opinion Spam Analysis. In the Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, CIKM'15, pages 1131-1140, 2015.

[7] V. Sandulescu, and M. Ester. Detecting Singleton Review Spammers Using Semantic Similarity. In the Proceedings of the 24th International Conference on World Wide Web, WWW'15, pages 971-976, 2015.

[8] A. Heydari, M. Tavakoli and N. Salim. Detection of fake opinions using time series. In Expert Systems with Applications, pages 83-92, 2016.

[9] Santosh KC, A. Mukherjee. On the Temporal Dynamics of Opinion Spamming: Case Studies on Yelp. In the Proceedings of the 25th International Conference on World Wide Web, WWW'16, pages 369-379, 2016.

[10] Siddu P. Algur , Jyoti G. Biradar, P. Bhat. Multidimensional Time Series Based Review Spam Detection. In the International Journal of Innovative Research in Computer and Communication Engineering, 2016.

[11] R. Mihalcea, C. Corley, and C. Strapparava. Corpus-based and knowledge-based measures of text semantic similarity. In the Proceedings of the 21st National Conference on Artificial intelligence, AAAI'06, pages 775-780, 2006.

[12] Quoc Le, Tomas Mikolov. Distributed representations of sentences and documents. In the Proceedings of the 31st International Conference on International Conference on Machine Learning, ICML'14, pages II-1188-II-1196, 2014.

[13] Luyang Li, Bing Qin, Wenjing Ren, Ting Liu. Document representation and feature combination for deceptive spam review detection. In Neurocomputing, pages 33-41, 2017.

[14] https://www.amazon.com/POR-15-MRQ-Metal-Ready/dp/B000ET60TM

[15] https://www.amazon.com/POR-15%C2%AE-Cleaner-Degreaser-Surface-Bottle/dp/B004LY5W2M

LIST OF PUBLICATIONS

---

[1] Khushbu Jhunjhunuwala and Durga Toshniwal, "Spam Detection using Semantic and Temporal Analysis in Reviews" in 28th International Joint Conference on Artificial Intelligence: The 1st International Workshop on Artificial Intelligence for Business Security.(ERA A, Qualis A1) *(Submitted)*