

SPAMMER DETECTION AND FAKE USER IDENTIFICATION ON SOCIAL MEDIA

SRIHARIHARAN T¹, SHANMUGASURYA S², SENTHILRAJA P³

¹srihariharant@gmail.com, Student, Department of Computer Science and Engineering, K.S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu

²shanmugasurya1999@gmail.com, Student, Department of Computer Science and Engineering, K.S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu

³senthilraja@ksrct.ac.in, Assistant Professor, Department of Computer Science and Engineering, K.S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu

ABSTRACT

As e-commerce is growing and becoming popular day-by-day, the number of reviews received from customer about any product grows rapidly. People nowadays heavily rely on reviews before buying anything. Product reviews play an important role in deciding the sale of a particular product on the ecommerce websites or applications like Flipkart, Amazon, Snapdeal, etc. In this paper, we propose a framework to detect fake product reviews or spam reviews by using Opinion Mining. The Opinion mining is also known as Sentiment Analysis. In sentiment analysis, we try to figure out the opinion of a customer through a piece of text. The proposed method called VWNB-FIUT (Value Weighted Naïve Bayes with Frequent Pattern Ultra Metric Tree) automatically classifies users' reviews into "suspicious", "clear" and "hazy" categories by phase-wise processing. The hazy category recursively eliminates elements into suspicious or clear. This results into richer detection and be useful to business organization as well as to customers. Business organization can monitor their product selling by analyzing and understanding what the customers are saying about products. This can help customers to purchase valuable product and spend their money on quality products. Finally end users see that each individual review with polarity scores and credibility score annotated on it. We first take the review and check if the review is related to the specific product with the help of VWNB. We use Spam dictionary to identify the spam words in the reviews by using FIUT. In Text Mining we apply several algorithms and on the basis of these algorithms we get the specific results.

1. INTRODUCTION

1.1 ECOMMRCE PLATFORM

An ecommerce platform is a software application that allows online businesses to manage their website, marketing, sales, and operations. E-commerce is the activity of buying or selling of products on online services or over the Internet. Electronic commerce draws on technologies such as mobile commerce, electronic funds transfer, supply chain management, Internet marketing, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data collection systems.

1.2 OPINION MINING

Product inference mining is a process of tracking the mood of the public about a particular product. Opinions can be essential when it's use to make a decision or choose among multiple option. Information-gathering behavior has always been to find out what other people think. The availability of opinion-rich resources such as online review sites and personal blogs, and challenges arise, to understand the opinions of others people.

1.2.1 LEVELS OF PRODUCT INFERENCE MINING

Product inference mining is extracting people's opinion from the web. It is also known as sentiment analysis. There are three tasks for opinion mining [14]

- Document-level opinion mining
- Sentence-level opinion mining
- Phrase-level opinion mining

1.2.2 TASKS IN OPINION MINING

The area of opinion analysis is to predict the polarity of a piece of opinion text as positive or negative. Product inference analysis tasks unnoticed due to lack of popularity. Here, the tasks related to opinion analysis are,^[4]

- Subjectivity Detection
- Sentiment Prediction
- Aspect Based Sentiment Summarization
- Contrastive Viewpoint Summarization
- Text Summarization for Opinions
- Predicting Helpfulness of Online Comments/Reviews
- product inference-Based Entity Ranking

1.3 SPAMMER GROUP

We focus on group spam, which has not been studied so far. A spammer group refers to a group of reviewers who works together writing fake reviews to promote or demote a set of target products. Spammer groups are very damaging due to their sheer sizes. It is well-known that many online reviews are not written by genuine users of products, but by spammers who write fake reviews to promote or demote some target products. Although some existing works have been done to detect fake reviews and individual spammers, to our knowledge, no work has been done on detecting spammer groups. This work focuses on this task and proposes an effective technique to detect such groups.

2. LITERATURE REVIEW

2.1 IMPACT OF ONLINE CONSUMER REVIEWS ON SALES

F. Zhu and X. Zhang et.al “Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics,” in this work how product and consumer characteristics moderate the influence of online consumer reviews on product sales using data from the video game industry. The findings indicate that online reviews are more influential for less popular games and games whose players have greater Internet experience.

The differential impact of consumer reviews across products in the same product category and suggests that firms’ online marketing strategies should be contingent on product and consumer characteristics ^[1].

2.2 TEMPORAL DYNAMICS OF OPINION SPAMMING

K. C. Santosh and A. Mukherjee et.al “on the temporal dynamics of opinion spamming: Case studies on yelp,” In this work recently, the problem of opinion spam has been widespread and has attracted a lot of research attention. While the problem has been approached on a variety of dimensions, the temporal dynamics in which opinion spamming operates is unclear.

How does buffered spamming operate for entities that need spamming to retain threshold popularity and reduced spamming for entities making better success? We analyze these questions in the light of time-series analysis on Yelp. This work analyses discover various temporal patterns and their relationships with the rate at which fake reviews are posted. Building on our analyses, we employ vector auto regression to predict the rate of deception across different spamming policies.

Next, we explore the effect of filtered reviews on (long-term and imminent) future rating and popularity prediction of entities. Our results discover novel temporal dynamics of spamming which are intuitive, arguable and also render confidence on Yelp’s filtering. Lastly, we leverage our discovered temporal patterns in deception detection. Experimental results on large-scale reviews show the effectiveness of our approach that significantly improves the existing approaches ^[2].

2.3 OPINION SPAM AND ANALYSIS

N. Jindal and B. Liu et.al, “Opinion spam and analysis,” in this work 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. 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. We exploit machine learning methods to identify review spam. Toward the end, we manually build a spam collection from our crawled reviews. We first analyze the effect of various features in spam identification. We also observe that the review spammer consistently writes spam.

This provides us another view to identify review spam: we can identify if the author of the review is spammer. Based on this observation, we provide a two view. Semi-supervised method, co-training, to exploit the large amount of unlabeled data. The experiment results show that our proposed method is effective. Our designed machine learning methods achieve significant improvements in comparison to the heuristic baselines [3].

3. PROPOSED SYSTEM

In this proposed system the Value Weighted Naïve Bayes with Frequent Pattern Ultra Metric Tree based opinion review analysis reviews possess the following characteristics: (a) they are frequently commented in user reviews; and (b) users' opinions on these reviews greatly influence their overall opinions on the reviews. A straightforward frequency-based solution is to regard the reviews that are frequently commented in user reviews as important. However, users' opinions on the frequent reviews may not influence their overall opinions on the reviews, and would not influence their purchasing decisions. We are measuring public concern using a two-step sentiment word alignment approach.

This work VWNB-FIUT Identifying fake reviews from a large dataset is challenging enough to become an important research problem. Business organizations, specialists and academics are battling to find the best system for opinion spam analysis. A single algorithm cannot solve all the problems and challenges faced in today's generation with advancements in technologies, though a few are very efficient in analysis. It also improving the performance of the opinion spam analysis, and developing one that is consistently efficient across all categories of data.

The opinion reviews obtained from users can be classified into positive or negative reviews, which can be used by a consumer to select a product. This work aims to classify amazon reviews into groups of positive or negative polarity by using machine learning algorithms. In this study, we analyze online amazon reviews using proposed methods in order to detect fake reviews. The text classification methods are applied to a dataset of Amazon or google play store reviews.

ADVANTAGES

- The reviews containing explicit content and with swear words are not taken into consideration and are removed from the dataset.
- Sentiment score for each word is calculated when words are extracted into a form of dictionary or so called 'Bag of Words (BOW)' It first identifies the nouns and noun phrases in the documents. The occurrence frequencies of the nouns and noun phrases are counted, and only the frequent ones are kept as reviews.
- The language model was built on reviews, and used to predict the related scores of the candidate reviews. The candidates with low scores were then filtered out.
- The admin can easily identify related opinion reviews on that session.
- Easily determine reviews quality by using customer reviews.
- We can find Based on the number of reviews classified as Personal Negative; we compute a Measure of Concern (MOC) and a timeline of the MOC. We attempt to correlate peaks of the MOC timeline to peaks of the News (Non-Personal) timeline.
- Best accuracy results are achieved.
- Analysis of product after spam removal is done on the basis of their respective features.

4. CONCLUSION AND FUTURE ENHANCEMENT

This work proposes a partially supervised learning-based Model VWNB to detect spammer groups from product reviews. First, the frequent item mining using the VWNB model (FIM) to discover spammer group candidates from the review data. Then, manually labeling some spammer groups as positive instances, the VWNB employs construct to PU-Learning the positive and unlabeled instances of a classifier to identify the real candidates from the group of real spammer groups. In particular, the VWNB dense a feature strength functions the group features of the measure of discriminatory power, and then High discriminative with iteratively removes instances Get a reliable set of unlabeled instances from only non-spammer groups of negative set consisting. By combining the positive, negative and unlabeled instances, we the well-known semi supervised into the PU-Learning problem learning problem, and employer Naive Bayesian Model and EM algorithm to construct a classified as spammer group detector. Experiments on Amazon.cn show that the proposed VWNB model outperforms both

supervised and Spammer group detection on unsupervised learning methods. Improvement in the area of our future work of the VWNB model.

5. RESULT AND DISCUSSION

The malignant exercises on online media are being acted in a few different ways. In addition, the scientists have endeavored to recognize spammers or spontaneous bloggers by proposing different arrangements. Accordingly, to consolidate every appropriate exertion, we proposed a scientific categorization as per the extraction and classification strategies. The classification depends on different classifications, for example, counterfeit content, URL based, moving points, and by recognizing counterfeit clients. The rest significant classification in the scientific categorization is of strategies proposed for recognizing spam, which is infused in the Twitter stage through phony substance. Spammers for the most part consolidate spam information with a point or catchphrases that are malignant or contain the sort of words that are probably going to be spam.

6. REFERENCES

- [1] F. Zhu and X. Zhang, "Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics," *J. Marketing*, vol. 74, no. 2, pp. 133–148, 2010.
- [2] K. C. Santosh and A. Mukherjee, "on the temporal dynamics of opinion spamming: Case studies on yelp," in *Proc. 25th Int. Conf. World Wide Web*, 2016, pp. 369–379.
- [3] N. Jindal and B. Liu, "Opinion spam and analysis," in *Proc. Int. Conf. Web Search Data Mining*, 2008, pp. 219–230.
- [4] F. Li, M. Huang, Y. Yang, and X. Zhu, "Learning to identify review spam," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2011, vol. 22, no. 3, pp. 219–230.
- [5] S. Xie, G. Wang, S. Lin, and P. S. Yu, "Review spam detection via temporal pattern discovery," in *Proc. 18th ACM SIGKDD Int. Conf. Know. Discovery Data Mining*, 2012, pp. 823–831.
- [6] A. Mukherjee et al., "Spotting opinion spammers using behavioral footprints," in *Proc. 19th ACM SIGKDD Int. Conf. Know. Discovery Data Mining*, 2013, pp. 632–640.
- [7] L. Akoglu, R. Chandy, and C. Faloutsos, "Opinion fraud detection in online reviews by network effects," in *Proc. ICWSM*, vol. 13, 2013, pp. 2–11.
- [8] S. Shehnepoor, M. Salehi, R. Farahbakhsh, and N. Crespi, "Net Spam: A network-based spam detection framework for reviews in online social media," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 7, pp. 1585–1595, Jul. 2017.
- [9] Z. Wu, Y. Wang, Y. Wang, J. Wu, J. Cao, and L. Zhang, "Spammers detection from product reviews: A hybrid model," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2015, pp. 1039–1044.
- [10] A. Heydari, M. Tavakoli, and N. Salim, "Detection of fake opinions using time series," *Expert Syst. Appl.*, vol. 58, pp. 83–92, Oct. 2016.
- [11] H. Li et al., "Bimodal distribution and co-bursting in review spam detection," in *Proc. 26th Int. Conf. World Wide Web*, 2017, pp. 1063–1072.
- [12] A. Mukherjee, B. Liu, and N. Glance, "Spotting fake reviewer groups in consumer reviews," in *Proc. 21st Int. Conf. World Wide Web*, 2012, pp. 191–200.
- [13] C. Xu, J. Zhang, K. Chang, and C. Long, "Uncovering collusive spammers in Chinese review websites," in *Proc. 22nd ACM Int. Conf. Conf. Inf. Know. Manage.* 2013, pp. 979–988.
- [14] Y. Wang, Z. Wu, Z. Bu, J. Cao, and D. Yang, "Discovering shilling groups in a real e-commerce platform," *Online Inf. Rev.*, vol. 40, no. 1, pp. 62–78, 2016.
- [15] B. Liu and W. S. Lee, "partially supervised learning," in *Web Data Mining*. Berlin, Germany: Springer, 2011, pp. 171–208.