

## A Study on: Opinion/Review Spam Detection

Sanketi P. Raut  
ME II (Computer),  
Lokmanya Tilak College of Engineering,  
New Mumbai,  
Email:sanketiraut28@gmail.com

Prof. Chitra Wasnik  
Computer Dept.  
Lokmanya Tilak College of Engineering,  
New Mumbai,  
Email:chitrawasnik@yahoo.com

**Abstract** - The most common mode for consumers to express their level of satisfaction with their purchases is through online ratings, which we can refer as Online Review System. Network analysis has recently gained a lot of attention because of the arrival and the increasing attractiveness of social sites, such as blogs, social networking applications, micro blogging, or customer review sites. Online review systems plays an important part in affecting consumers' actions and decision making, and therefore attracting many spammers to insert fake feedback or reviews in order to manipulate review content and ratings. Malicious users misuse the review website and post untrustworthy, low quality, or sometimes fake opinions, which are referred as Spam Reviews.

In this study, we aim at providing an efficient method to identify spam reviews and to filter out the spam content.

**Keywords**- Review Spam detection, Opinion, Text mining, WordNet, Language model, Tree-base decision.

\*\*\*\*\*

### 1. INTRODUCTION

Every consumer needs secure trustworthy, reliable and easily available information about any product while making purchase. Opinions/Reviews are middle of almost all human actions and are main influencers of our behaviours. Review system plays an important role in decision making as many People use the reviews while deciding quality of product to purchase the top quality product. Companies or merchants use opinions or feedback to take a decision to improve their sales according to intellectual things done by other challengers. This motivates some corrupt merchants to put on fake reviews in order to mislead the prospective consumers by improving their reputation or shrinking the competitors'. Thus, it brings a crucial demand to detect fake reviews in early stage for reducing their influence. The Existing system makes use of social relationships between users of system. It is based on the identification of review spammers by incorporating social relations based on two assumptions that people are more likely to consider reviews from those connected with them as trustworthy, and review spammers are less likely to maintain a large relationship network with normal users and shows a strong correlation between social Relationships and the overall trustworthiness scores [1]. The standard collaborative filtering (CF) is not suitable for the new users who do not have much review history. The Existent method is not attack resistant, as fake reviews on system can affect the rating score and resulting in the misleading to other users. This paper proposes a method that will be identifying the fake review by first, analysing the review content, detecting whether the content is spam or not a spam and the providing a spam free result. The crucial goal of opinion spam detection in the review framework is to identify every fake review and fake reviewer.

### 2. MOTIVATION

We accept as truth that as opinions on the Web are gradually more used in practice by consumers, administrations, and companies for their decision making. These reviews or

opinions are advantageous only if the reviews posted appropriately without any incorrect intention. Day by day Opinion spamming is getting worse and also more sophisticated. Detecting spam reviews or opinions have become more and more critical. The position is at present quite bad. According to our annual Local Consumer Review Survey, the importance of online reviews in the purchasing decision is growing day by day.

Eighty-five percent of consumers are satisfied once they have read up to 10 online reviews for local businesses and to make purchases. Many customers even say that they wouldn't buy product without reading online reviews. Reviews become even more important in the situation where customers are not able to test products practically prior to purchasing it. But these reviews have some drawbacks which give motivation to review spam detection- bad publicity, some reviews are fake given for supporting or devaluing product.

When intention to give an opinion is not true, such opinion can be spam. Opinion spamming has become a most important issue. Apart from persons who give fake feedback in reviews and discussions, there are also business-related companies that are in the business of writing fake reviews and bogus information for their clients. There is need to detect such spamming activities to make sure that the Opinions/Reviews on the Web are a trustworthy source of important information. Therefore there is need to develop a system that will be helpful for managing reputation for business and avoiding people from getting misled.

### 3. OBJECTIVE

To propose an approach that will identify "illegal" activities (e.g., writing fake reviews/ Opinions) that try to mislead readers and to Propose a system that will filter out such spam content from reviews and prevent users from getting misled.

The main challenge of opinion spam detection is that dissimilar to other forms of spam, it is very hard to recognize fake feedback by manually reading them. It makes

the process of finding such fake data quite difficult. There is no mode to detect such fake review without considering information beyond the review text itself simply because the same opinion cannot be together truthful and fake at the same time.

#### 4. LITERATURE SURVEY

In [2], the review spam detection is concerned with a problem of singleton review, i.e. the reviewer written only one review using time series pattern discovery, in that they find the relationship between rating and volume of singleton reviews because as the review rises, rating also rises or drops dramatically. This paper gives a hierarchical framework for robust Singleton review spam detection. States that multi-scale anomaly detection algorithm on multidimensional time series is effective in detecting Singleton review spams. In [3], they proposes the technique for untruthful review spam detection that done using text mining model and integrated into semantic language model.

Non review spam detection is done by identifying different stylistic, syntactical and lexical features and SVM classifier applied to them. The results say that semantic language modelling and a text mining-based computational model are effective for the detection of untruthful reviews, even if spammers exercise confusion strategies.

In [4], proposes different behavior models which are based on review pattern. They focus on pattern of review content and ratings to define models, i.e. Targeting Product (TP), Targeting Group (TG), General Rating Deviation (GRD), and Early Rating Deviation (ERD) and finally shows that spammer has more impact on rating. In [5] they uses supervised learning with manually labelled training examples, detection of such spam is done by detecting duplicate reviews, to build a classification model only positive training examples are used. User feedback on reviews is not effective in filtering out spam.

#### 5. OPINION

An opinion contains of two components: a target  $t$  and a sentiment  $s$  on the target, i.e.,  $(t, s)$ , where  $t$  can be any entity or part of the entity for which an opinion has been expressed, and  $s$  is a positive, negative, or neutral attitude, or a numeric rating score expressing the strength/intensity of the sentiment (e.g. star ratings). Positive, negative, and neutral are called opinion orientations. An entity about which an opinion has placed can be any product, service, topic, issue, person, organization, or any event.

There are several types of opinion:

Direct opinion:

A direct opinion is an opinion spoken directly on an entity or an entity aspect,

e.g., "The cloth quality is great."

Indirect opinion:

An indirect opinion is an opinion that is expressed indirectly on an entity or aspect of an entity based on its effects on some other entities. This sub-type often occurs in the medical domain.

e.g., "After injection of the drug, my joints felt worse"

This describes a detrimental effect of the drug on "my joints", which not directly gives a negative opinion or sentiment to the drug.

Comparative opinion:

A comparative opinion is an opinion that expresses a relation of similarities or differences between two or more entities.

e.g., "Coke tastes better than Pepsi"

#### 6. PRAPOSED SYSTEM

The propose system divide work flow as

- a. It Performs Sentiment classification that determines whether a Review is positive, negative or neural.
- b. Featured base-opinion mining that discovers features of a reviewed entity with the intent of acquiring the opinion of a reviewer about that specific feature and providing a spam free content.

The propose work discuss modules that will perform the process as Review spam detection, non-Review spam detection , Brand spam detection and filter out the spam content.

##### i. Query Input:

Provide Review Query as an input to the system in order to extract the review content. The text of a review is called the content of the review. The content of each review is the first thing to be considered while detection of spam. Linguistic feature such as word and POS  $n$ -grams can be used for distinguishing malicious behaviours (deceptions and lies) which can be extracted from content of a review.

##### ii. Analysing Customer Review:

Analysis of customer Review is needed to produce many spam features. This phase identifies two things. First, it Identifies Meta-data of Review, the Information about the review besides its actual content is called meta-data. Because sometimes it can found that several user-ids write multiple positive or negative reviews of a particular entity using the same system or it can happen that specific reviewer has posted a number of positive reviews for products of a particular brand and many negative reviews for products of other competing brands.

For e.g., the reviewer's identity, the time of the review's publication, the star rating given by reviewer, the time duration of writing the review and its MAC and IP addresses etc.

Through analysing these types of data, system can identify certain malicious behaviours of reviewer. Second, it identifies information about product about which the Review has been placed. Information about a product is very useful in spam detection. Such information can Helps to identify whether the review is about the same product??

A feature extraction process must be performed on text documents before they can be mined. This pre-processing step transforms text documents into small units of text which are called as features of the product. Feature selection process contains feature identification and feature cropping. Feature selection plays vital role while mining review content. Quantified and unquantified features are brought together to get the better performance.

### iii. Text mining:

Text mining also referred as text data mining, roughly equal to text analytics, it is a process of deriving high-quality information from the given text. Text mining is the analysis of data contained in natural language text, mining unstructured data with natural language processing, collecting or identifying a set of textual materials. Text mining will follow text categorization (TC), which performs classification of Review text with respect to a set of one or more preexisting categories.

Text categorization is used for Finding Text duplication and conceptual similarity between reviews. This explores a method that uses WordNet concept for categorizing the Review Text. WordNet is a thesaurus for the English language based on psycholinguistics studies. WordNet ontology will capture the relations between the Review words. It refer as a data-processing resource which covers lexico-semantic categories called synsets (Synonym Sets) [6]. The synsets are sets of synonyms which gather lexical items having similar significances. I.e. certain adjectival concepts which meaning is close are gathered together. The hyponymy is represented in WordNet is interpreted by "is-a" or "is a kind of" relationships.

### iv. Non-Review Spam detection:

Such type of reviews includes no Opinion. It may contain advertising about variety of products, sellers, other irrelevant things such as questions, answers or similes, some random text etc. To identify and filter out such spam content classifier is useful. Such content are considered as spam because they are not giving any opinion.

### v. Untruthful Review Spam Detection:

These are reviews that are written not based on the reviewers' genuine experiences of using the products or services, but are written with some secret intentions.

In such type of Review, the reviewer often post more positive or more negative review about some product. While finding Untruthful reviews input to the system I s set of all reviews about same product, calculate the probability of word sequences of review. Set the Threshold value, and the probability is used to decide whether review is positive or negative. More Positive reviews are the opinion expressing a worthless positive feedback of a product with the intension of promoting that product. More Negative reviews are the opinion, expressing a spiteful negative feedback about a product with the intension of damaging status of product. Language model is useful while detecting and filtering out the spam content.

### vi. Brand Spam detection:

Such reviews are not posted on any product, instead it is posted on specific brand, manufacturer or seller of product. To find this spam, there is a need to find features used in Reviews, these features helps to determine whether the review is about the product or about any brand. Feature selection decision tree algorithm is useful for finding the feature from reviews that do not give wrong interpretation of phrases in the reviews. Then for each possible combinations of features from review text, check for the root to leaf node of decision tree, by calculating frequency of low referenced features, remove it. Then according to the features gained

and features mentioned in tree, decision is taken whether the review is on brand or not

## 7. EXPECTED RESULT

At end, result of spam detection is analyzed and decision will be taken on -whether each review is spam or not a spam. Such result is helpful to both users and vendor application during making their respective decisions. System will be giving Spam free Result.

As individual users and companies use reviews and opinion for decision making, it is important to detect opinion spam and opinion spammer. This approach mainly concentrates on non-review spam detection, untruthful review spam detection and brand spam detection and filtering. The result will give more accuracy to display a Spam free Review which is helpful to both customers while buying any product and for company to improve their performance using true reviews.

## 8. CONCLUSION

The recent work related to spam detection is done with classifier, language model and Decision tree, which gives more efficiency and trustworthiness while detecting and filtering the spam content. The results are promising. Supervisors, controllers, organizations can use review spam detection result as an administrative tool to supervise target e-commerce accumulation. The system gives convenience to administrators, flexible settings are available. It provides efficient and trustworthy opinion and feedback.

## REFERENCES

- [1] Hao Xue\*, Fengjun Li\*, Hyunjin Seo†, and Roseann Pluretti , "Trust-Aware Review Spam Detection", EECS Department, †School of Journalism and Mass Communications The University of Kansas Lawrence, KS, USA, 2015 IEEE.
- [2] Sihong Xie, Guan Wang, Shuyang Lin, Philip S. Yu "Review spam detection via time series pattern discovery", ACM Proceedings of the 21st international conference companion on World Wide Web, pp.635-636, 2012.
- [3] Raymond Y. K. Lau, S. Y. Liao, Ron Chi-Wai Kwok, Kaiquan Xu, Yunqing Xia, Yuefeng Li, "Text mining and probabilistic modeling for online review spam detection" ACM Transactions on Management Information Systems (TMIS) , Volume 2 Issue 4, Article 25, 2011.
- [4] Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal, Bing Liu, Hady Wirawan Lauw , "Detecting product review spammer using rating behaviors", Proceedings of the 19th ACM international conference on Information and knowledge management, pp-939-948, 2010.
- [5] Nitin Jindal and Bing Liu "Analyzing and Detecting Review Spam" *Department of Computer Science University of Illinois at Chicago* nitin.jindal@gmail.com Seventh IEEE International Conference on Data Mining, © 2007 IEEE.
- [6] "Using WordNet for Text Categorization" Zakaria Elberrichi1 , Abdelattif Rahmoun2 , and Mohamed Amine Bentaalal1 1 EEDIS Laboratory, Department of Computer Science, University Djilali Liabès, Algeria 2 King Faisal University, Saudi Arabia
- [7] [www.wikipedia.org](http://www.wikipedia.org)