Fake Product Review Monitoring System

B. Suganthi and Ms. K. Manohari (MCA, M.Phil),
[1,2]Deivannaiammal College for Women, Villupuram, Tamilnadu, India

Abstract: In the current scenario, the data on the web is growing exponentially. Social media is generating a large amount of data such as reviews, comments, and customer's opinions on a daily basis. This huge amount of user generated data is worthless unless some mining operations are applied to it. As there are a number of fake reviews so opinion mining technique should incorporate Spam detection to produce a genuine opinion. Nowadays, there are a number of people using social media opinions to create their call on shopping for product or service. Opinion Spam detection is an exhausting and hard problem as there are many faux or fake reviews that have been created by organizations or by the people for various purposes. They write fake reviews to mislead readers or automated detection system by promoting or demoting target products to promote them or to degrade their reputations. The proposed technique includes Ontology, Geo location and IP address tracking, Spam words Dictionary using Naïve Bayes, Brand only review detection and tracking account used.

I. INTRODUCTION

One of the very rapid growth area is ecommerce. Generally e-commerce provide 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 but 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. They share their thoughts on internet.

Before purchasing anything, it is a normal human behavior to do a survey on that product. Based on reviews, customers can compare different brands and can finalize a product of their interest. These online reviews can change the opinion of a customer about the product. If these reviews are true, then this can help the users to select proper product that satisfy their requirements. On the other hand, if the reviews are manipulated or not true then this can mislead user. This boosts us to develop a system which detect fake reviews for a product by using the text and rating property from a review. The honesty value and measure of a fake review will be measured by utilizing the data mining techniques.

An algorithm could be used to track customer reviews, through mining topics and sentiment orientation from online customer reviews and will also blocked the fake reviews.

II. RELATED WORK

A. Opinion Mining Using Ontological Spam Detection

Duhan & Mittal proposed a paper "Opinion Mining Using Ontological Spam Detection" which will help us to find out fake reviews by using Naïve Bayes as algorithm. To find out fake review in the website this "Fake Product Review Monitoring System" system is introduced. This system will find out fake reviews made by the customers and it will block the users. To find out the review is fake or genuine, we will use some classification such as

- **Tracking IP address** of the user to detect if the reviews are from a Spammer. If multiple reviews are from the same IP address then the Reviews are considered Spam.
- **Using Account Used** to check whether the reviews are done using the same account.
- **Brand only Review detection** i.e. whether the reviews are on only Brand not the product. It's not helpful to consider only the Brand value to judge a product.
- **Using Negative Dictionary** i.e. the negative words are identified in the review. If there are more than five Negative Words then the review is a Spam.

For instance, a user has posted a Review: "This product is not good, the design is bad, quality is worst and it is worthless to buy." Here, this sentence consists of 4-5 negative words. So, the system will check the count of negative words, if the count exceeds, then it will be considered as spam review. Therefore Negative Word

Dictionary will be used with customized Senti strength algorithm. According to this approach, probability of given review to be Spam is more so it will be considered a Spam.

- **Using Ontology:** For instance, if the review posted on a product is not about that product but talking about something else then ontology is used to identify and classify such reviews as spam.

If Class: Toshiba

Context: Laptop Review: *Dell is not so good.*

Here User is Posting Reviews about Laptop that comes under the class Toshiba. But his Review contains Dell Keyword. In order to identify this Review as Spam we are going to use Ontology.

This system uses data mining methodology and Opinion mining technology. This system helps the user to find out correct review of the product, will also help the user to detect fake review and makes them to block the fake reviews automatically.

B. Fake Product Review Monitoring and Removal for Genuine Online Product Reviews Using Opinion Mining

Kohli, Mishra & Gupta proposed a paper "Fake Product Review Monitoring and Removal for Genuine Online Product Reviews Using Opinion Mining" which help us in detecting the fake reviews and track down the user. As most of the people require review about a product before spending their money on the product. So people come across various reviews in the website but these reviews are genuine or fake is not identified by the user. In some review websites some good reviews are added by the product company people itself in order to make product famous this people belong to Social Media Optimization team.

They give good reviews for many different products manufactured by their own firm. User will not be able to find out whether the review is genuine or fake. To find out fake review in the website this "Fake Product Review Monitoring

and Removal for Genuine Online Product Reviews Using Opinion Mining" system is introduced. This system will find out fake reviews made by the social media optimization team by identifying the IP address. User will login to the system using his user id and password and will view various products and will give review about the product. And the user will get genuine reviews about product. And while reviewing he needs to enter the email id from which he is reviewing and it would be verified. If he writes a fake review then his id will be blocked bot allowing him to share his opinions again.

System works as follows:-

- Admin will add products to the system.
- User need to enter their email id and OTP no to enter the system
- User once access the system, user can view product and can post review about the product.
- For posting reviews, the user's id will be verified.
- And admin will also block the email id of the user if reviews are spammed.
- Admin will delete the review which is fake.
- Admin Login: - Admin login to the system using his admin ID and password.
- Add product: - Admin will add product to the system.
- Delete Review: - Admin will remove the review which tracked by the system as fake.
- User Login: - User will login to the system using his user ID and password.
- View product: - User will view product.
- Post Review: - User can post review about the product.

III. PROBLEM DEFINITION AND OBJECTIVES

A. Problem Definition

In recent years, online reviews have been playing an important role in making purchase decisions. This is because, these reviews can provide customers with large amounts of useful information about the goods or service. However, to promote factitiously or lower the quality of the products or services, spammers may forge and produce fake reviews. Due to such behavior of the spammers, customers would be mislead and make wrong decisions. Thus detecting fake (spam) reviews is a significant problem. Opinion spamming refers to the use of excessive and illicit methods such as creating a large volume of fake reviews, in order to generate biased positive or negative opinions for a target product or service with the intention of promoting or demoting it, respectively. The reviews created for this purpose are known as fake, spam or bogus reviews, and the authors responsible for composing such deceptive content are known as fake or spam reviewers.
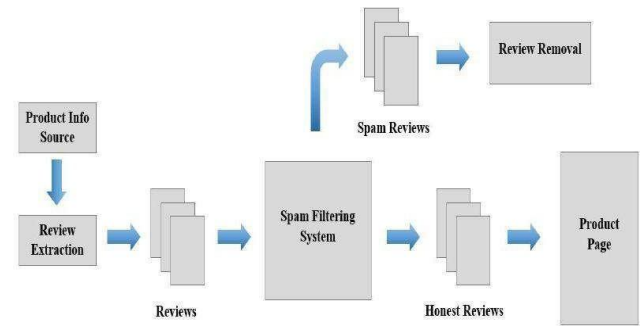
B. Objectives

The identified challenges motivate to bring up a solution to all the problems stated in the above problem statement section. Following are the objectives of the proposed approach and this thesis work:

- To implement different algorithm to get better Spam Detection i.e.; IP Address, Account used, Negative Word Dictionary using Senti-strength, Ontology.
- Graphical representation of work.
- To deals with 6 different types of Spam Reviews.
- To presents Opinion Mining on Spam Filtered Data.
- To implement Ontology in Spam Detection

- To present an algorithm that does Opinion Mining with Spam Detection.

IV. FLOWCHART

SUMMARY

They are various ways to detect Spam Reviews in order to the Opinion mining to be more accurate and useful have been studied. A detailed discussion about the existing techniques, to find out the whether the review is spam or not is presented. Other Techniques are incorporated like IP Address Tracking and Ontology to detect Spam Reviews in order to get more accurate results from Opinion mining.

After detecting the spam reviews from the existing Dataset, a new Dataset is created which doesn't contain spam reviews and then opinion mining is performed on the new Spam Filtered Dataset. At last a new algorithm is proposed that detects spam reviews more precisely and performs opinion mining using spam filtered data.

Working

1. User will be allowed to review only if he is logged into our online portal.

2. After logging in user will be allowed to review for the product.

3. Once the user enters the review, the reviews will be processed and analyzed for spam on following conditions:

a) Does the review entered by the user contain any link which redirects them to other product page for brand promotion?

b) Analyzing whether multiple review have come from the same user.

c) Analyze whether same email account or same ip- address are used for multiple reviews on same product.

d) Analyze the reviews or ratings to detect whether reviews are spam or not.

4. If the review posted by the user satisfies any of the above specified conditions then it will be considered as spam/fake reviews.

5. Once the review is detected as spam review or fake review, then user account will be blocked and review will be reported to the administrator.

References

[1] Rajashree S. Jadhav, Prof. Deipali V. Gore, "A New Approach for Identifying Manipulated Online Reviews using Decision Tree ". (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2), pp 1447-1450, 2014

[2] Long- Sheng Chen, Jui-Yu Lin, "A study on Review Manipulation Classification using Decision Tree", Kuala Lumpur, Malaysia, pp 3-5, IEEE conference

publication, 2013.

[3] Benjamin Snyder and Regina Brazil, "Multiple Aspect ranking using the Good Grief Algorithm "Computer Science and Artificial Intelligence Laboratory Massachusetts Institute of Technology2007.

[4] Ivan Tetovo, "A Joint Model of Text and Aspect Ratings for Sentiment Summarization "Ivan Department of Computer Science University of Illinois at Urbana, 2011

[5] N. Jindal and B. Liu, "Analyzing and detecting review spam," International Conference on Web Search and Data Mining, 2007, pp. 547-552.

[6] N. Jindal and B. Liu, "Opinion spam and analysis," International Conference on Web Search and Data Mining, 2008, pp. 219-230.