

Detecting Fraudulent Manipulation of Online reviews

Shwetha Yadav¹, Dr. Deven Shah²

¹Shwetha Yadav, PG student, Thakur College of Engineering and Technology, Mumbai, India

²Dr. Deven Shah, Professor, Dept of Information Technology, Thakur College of Engineering and Technology, Mumbai, India

Abstract - Reviews are basically feedback given by the customers on a particular product. Products can be of any type, for example, electronics (mobile phones, laptops, computers, etc.), home and kitchen products (utensils, refrigerator, etc.), dresses, cars, houses and many more. And the reviews given by the customers can be positive or negative.

People check the reviews before buying the product online. And hence the reviews must be trusted and true. Trust is an important factor in any e-commerce website. In the past few decades, the trading was done face-to-face. The customer can see the seller and even the product. Verify the product's quality, negotiate and then finally decides to buy. Thus it was possible, that both the customers and the sellers could have trustworthiness on each other.

However, in E-commerce website, there is a lack of this kind of direct trust. Although there are many technologies like cryptography, digital signature, certificates and many more helps the customers to trust blindly on the E-commerce website for the transaction purpose. But, one cannot trust the reviews because reviews can fake or fraud. These fake reviews are published either to sell maximum products by the organization itself or by the competitors to degrade the fame and reputation of the organization or by the customers. And hence it is important to detect fraudulent manipulation of online reviews.

In this project, a new architecture is proposed for TRS in an E-commerce application. In which it includes feedbacks' analysis. A proposed system calculates the trustworthiness of the user and blocked the user who is trying to publish the fake reviews. It helps the user to compare the two products

Key Words: TRS, E-commerce, reviews, sentiments, Trust

1. INTRODUCTION

1.1. E-Commerce Website:

Before going to the topic of what is online reviews, why is it important? and why is there a need to detect the fake reviews? Let us start with the basic concept of what is an e-commerce website and online shopping.

E-commerce is nothing but buying the product and selling the product on the same platform which is an online mode. Seller sell the product online and the buyer buys the product

online. If we look in the past from where this all started, it was 11th August 1994 the boy named Phil Brandenberger from Philadelphia. He logged to his computer around noon that particular day and bought Sting's "Ten Summoners' Tales" by using his credit card for \$12.48 plus shipping.

E-commerce is electronic commerce, it is used for buying and selling the goods and services online and the transaction is done online. These business transaction occurs either B2B (Business to Business), B2C (Business to Consumer), C2C (Consumer to Consumer), and C2B (Consumer to Business). Some of the examples are Amazon (B2C), Olx (C2C), IBM (B2B), Blogs (C2B).

In today's digital world, most of the consumers prefer e-commerce, because of the lucrative offers, but primarily because they have a review and feedback system to judge the product. It is also now a common practice amongst customers to post reviews about a product they purchase, be it positive or negative. They are also used by product manufacturers to identify strengths and problems in their products and to find competitive intelligence, such as its potential worth in the market.

Nowadays, people like to shop online rather go outside, travel to the shop and then buy the required product. It is time-consuming and even energy consuming. People buy anything right from the basic necessities that are food, clothing, and shelter to the other stuff like ornaments, mobile phones, cars, or computers, to make life easy and comfortable.

So, there is need to know everything of online shopping, how it works, are they selling the right product, what's the quality of the product, whether it's good or bad, and many more. And from where we get this information? Yes, the information of each product's quality, it's actual usage, gets by the customer. The customers write a review of the product after buying it. They give feedback on a particular product. And these reviews are helpful for the other customer who is willing to buy the product, and it's a decision whether to buy or not.

1.2. Why Online reviews are essential:

- **Sale increases:**

If the reviews are positive and in large quantity then it helps in promoting the brand and website. People

obviously buy the product from your website, if they find the reviews are positive.

- **E-commerce website can be trusted:**
As user buys the product after seeing online reviews they will find the product accordingly, and hence the e-commerce site will be trusted for the services or the product they provide.
- **Impact on decision making:**
Customer's entire decision is based on the reviews which are published on the e-commerce website. So online reviews play an important role in decision making.

1.3. What are Fake reviews:

The fake review is nothing but the reviews which is not true and cannot be trusted. But these reviews are hard to find. Some fake reviews are written by the robots and might be identified by the organization but, some are written by the people itself and cannot be identified. And the customers cannot discriminate the fake reviews with the genuine one.

Fake reviews can be written by the organization itself to promote the products and sell the product to the customers in a large amount. It can also be written by the competitors to defame the organization's reputation. Some organizations can either attempt to help their deals by making a positive brand picture falsely or by creating counterfeit negative audits about a contender. The inspiration is, obviously, cash: online reviews are a huge business for movement goals, lodgings, specialist co-ops, and buy items.

As most people require reviews about a product before spending their money on the product. So people come across various reviews on the website but whether these reviews are genuine or fake, it cannot be identified by the user.

In some websites, the reviews are added by the company itself in order to make the product famous. So that user can buy the product and the organization can make a profit through it. User cannot discriminate whether the product is fake or the genuine one.

The fake reviews can be added by the competitors of the organization to defame the company's reputation or their product. This is harmful for the organization and the customers, therefore there is a need to detect fake reviews.

2. MOTIVATION

Product online reviews provide an information base for consumer experience. Online reviews appear on websites that sell products and offer services. Consumers are likely to look for information online before making purchase decisions.

Consumers will read online reviews, before deciding on the vacation destination. The quality of individual hotels is measured through different means, such as stars or ratings on characteristics of cleanliness, view, and staff friendliness. However, online reviews provide support for those means. It is critical to be able to identify fake online reviews with great accuracy.

There is a need to detect fake reviews because nowadays, before buying any product customers check the reviews of that particular product. And then decide whether to buy or not. To have trust on the e-commerce website, fake reviews should be detected and removed.

3. RELATED WORK

Table -1: Sample Table format

S r N o.	TITLE	AUTHORS	Y E A R	DISCRIPTION	GAPS
1.	Identificatio n of Fake Reviews Using Semantic and Behavioral Features	Xinyue Wang, Xianguo Zhang, Chengzhi Jiang, Haihang Liu	2018	Two types of features, one is behavioral feature set, the other is related to semantic.	Need to improve the existed algorithm which can be used to detect fake reviews more specifically and effectively.
2.	Review Spam Detection using Machine Learning	Dražko Radovanić	2018	A brief overview of spam detection methods published during the last decade was presented. It was shown that using different datasets yields extremely different results.	It should be focused on combining content-based and reviewer-based methods for achieving the best results.

3.	Fake Reviews Detection Based on LDA	Shaohua Jia, Xianguo Zhang, Xinyue Wang, Yang Liu	2018	Our study shows that although behavioral features reported very high detection accuracy with supervised classifiers, semantic features are also work well on real-life fake reviews in the commercial setting of Yelp.com.	Need of a high accuracy.
4.	Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques	Elshrif Elmurngi, Abdelouahed Gherbi	2018	We proposed several methods to analyze a dataset of movie reviews. Our research approaches studied the accuracy of all sentiment classification algorithms, and how to determine which algorithm is more accurate.	Focused only on movie reviews. Must have done research on many aspects.
5.	A Comparative Study on Fake Review Detection Techniques.	A.Lakshmi Holla, Dr Kavitha K.S	2018	Discusses the various techniques used for identifying fake reviews. Here an overview of the existing fake review detection methods along with their advantages and limitations are being clearly discussed.	There is a strong need to address the challenges in fake review detection and design suitable algorithms to improve the accuracy in detection.

4. PROPOSED WORK

In this project, we will calculate the trust degree of the user and the feedback's trustworthiness and generates the trust reputation score of the product. The program is used to determine fake reviews among the genuine ones.



Figure 1: System Flow.

Step 1: User writes the review:

Customer adds the review after buying the product or before buying it. The reviews which are added cannot be trusted because the reviews might be fake or genuine. Those reviews can harm the customers and as well as the organization. Reviews might be added to harm the reputation of the organization or else increase the sales of the product. It can be added by the competitors or by the organization itself depends on the situation. The customers might the defective product when they buy the product by checking the reviews.

Step 2: Algorithm used to detect Fake reviews from Genuine:

In this system the following steps will be carried out:

- A dummy e-commerce website will be created where the user can put the ratings/comments about the products.
- A knowledge-based database will store all the comments & ratings by the users.
- The main module will analyze the comments with the previously stored comments, compare them and the Trust and reputation systems (TRS) algorithm will commute the TRUST FACTOR for each user.
- The historical feedbacks by the users will also play a vital role in making the customer trustworthy.
- When both the conditions are satisfied, then the user's comment will be accepted by the site. If the user fails in the trust factor issue, the comment will not be accepted.
- User rating confidence/trustworthiness will be measured by based on the observations.

- The sentimental analysis will be carried out on the textual feedbacks for finding the user ratings and written feedback.
- The implementation will be helpful for the Production Company as well as vendors to decide trustworthy customers as well as the most popular products, which will surely help them in designing their future business policies.

How to calculate the trust score or trustworthiness of the user:

- It generates the user's trust degree which helps to calculate the trust score of the user's review.
- Indeed, each review has trustworthiness in a threshold [-5,5]. The most trustworthy the user's review is, the value of the trustworthy score is closer to 5. And if the review is untrustworthy then the value of the trustworthy score is closer to -5.
- If the review is trustworthy then the score of the user's review will be between [0,5], else it would be between [-5,0].

5. RESULTS AND DISCUSSIONS

The proposed architecture is helpful for detecting fraudulent manipulation of online reviews. It also helps the user to compare two products based on the reviews given by the different users. Few snapshots are attached bellow, to make understand the concept.

Features	Positive Polarity	Negative Polarity
Battery	2	0
Keyboard	0	2
HD	2	0
music	1	0
phone	14	16
processor	1	0
screen	2	0

Figure 2. Snapshot of sentiment score

The above figure is of sentiment score which is calculated using different reviews given by different users. It shows the score of the positive polarity and the negative polarity of different features.



Figure 3: Snapshot of calculated trust score

The above diagram is to calculate the trust score of the particular user. As you can see the score given by the system is 3.



Figure 4: Snapshot of temporary blocking the fake user.

The above snapshot shows that the user is blocked temporary because the system detects that the user is not trustworthy. So the user is blocked by the system temporarily. And his fake will not published by the system.

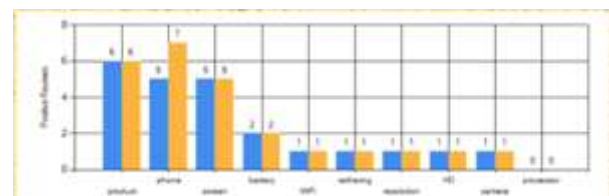


Figure 5: Comparing 2 products

The above figure shows the comparison of two products. There is an option to select the products and compare them.

6. CONCLUSION

The proposed Detection of Fraudulent manipulation of online reviews system helps to identify the fake reviewer and helps to block that user temporary. This system helps to publish only genuine reviews and not fake or manipulated reviews. In this project, a new architecture is proposed for TRS in an E-commerce application. In which it includes feedbacks' analysis. A proposed system calculates the trustworthiness of the user and blocked the user who is trying to publish the fake reviews. It helps the user to compare the two products. In this study, while comparing two product the algorithm takes time to implement, so in

future work, such an algorithm should be used which will reduce the time of implementation.

ACKNOWLEDGEMENT

I would like to express my gratitude to my Guide Dr. Deven Shah, Vice Pricipal, Professor, IT Department, TCET and ME-IT Coordinator Dr.Bijith Marakarakandy, Associate Professor, IT Department, TCET for their invaluable support and guidance throughout my P.G. research work. Without thier kind guidance & support this was not possible. I am grateful to them for timely feedback which helped me track and schedule the process effectively. They helped me to complete my project efficiently.

I would like to thank Dr. Kamal Shah, Dean, Professor, IT Department for her guidance and encouragement throughout my Post Graduation.

I would also like to thank Dr. B. K. Mishra, Principal, Thakur College of Engineering and Technology, for his encouragement and for providing an outstanding academic environment, also for providing the adequate facilities. I am thankful to all my M.E. teachers for providing advice and valuable guidance.

I also extend my sincere thanks to all the faculty members and the non-teaching staff and friends for their cooperation. Last but not the least, I am thankful to all my family members whose constant support and encouragement in every aspect helped me to complete my project.

REFERENCES

[1] D. Kornack and P. Rakic, "Cell Proliferation without 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] Nitin Jindal, Bing Liu, "Review Spam Detection", ACM proceedings of 16th international conference on world wide web, Pune, pp-1189- 1190, 2007.

[3] Ohana B and Tierney B, " Sentiment classification of reviews using SentiWordNet" 9th. IT & T Conference. 2009

[4] Li N and Wu D D, " Using text mining and sentiment analysis for online forums hotspot detection and forecast" Decision Support Systems, 2010, 48(2): 354-368

[5] Mudinas A, Zhang Dand Levene M, "Combining lexicon and learning based approaches for concept-level sentiment analysis", Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining. ACM, 2012: 5.

[6] Jindal Nitin, LIU Bing. "Review spam detection", Proceedings of the 16th international conference on World Wide Web. Canada. New York: ACM Press ,2007: 1189- 1190

[7] Mukherjee A, Venkataraman V, Liu B, Glance NS (2013) What yelp fake review filter might be doing? Boston, In ICWSM.

[8] N Jindal, B Liu, Proceedings of the 16th international conference on World Wide Web, 1189-1190

[9] N. Jindal and B. Liu. Analyzing and Detecting Review Spam. ICDM2007.

[10] A. Z. Broder. On the resemblance and containment documents. In Proceedings of Compression and Complexity of Sequences 1997, IEEE Computer Society, 1997

[11] Algur et al. (2010). Conceptual level similarity measure based review spam detection. In International conference on signal and image processing (pp. 8)

[12] Lai et al. (2010). Toward a language modeling approach for consumer review spam detection. In IEEE 7th international conference (pp. 8) [13] Pennebaker et al. (2007). The development and psychometric properties of LIWC2007. Austin, TX, LIWC.Net.

[14] Ott et al. (2011). Finding deceptive opinion spam by any stretch of the imagination. In 49th annual meeting of the association for the computational linguistics (pp. 11).

[15]Identification of Fake Reviews Using Semantic and Behavioral Features 2018 4th IEEE International Conference on Information Management Xinyue Wang, Xianguo Zhang, Chengzhi Jiang, Haihang Liu

[16] Fake Reviews Detection Based on LDA 2018 4th IEEE International Conference on Information Management, Shaohua Jia Xianguo Zhang, Xinyue Wang, Yang Liu.

[17] International Journal on Advances in Systems and Measurements, vol 11 no 1 & 2, year 2018, http://www.iariajournals.org/systems_and_measurements, Elsharif Elmurngi, Abdelouahed Gherbi.

[18] Identification of Fake Reviews Using Semantic and Behavioral Features 2018 4th IEEE International

Conference on Information Management, Xinyue Wang, Xianguo Zhang, Chengzhi Jiang, Haihang Liu.

[19] Review Spam Detection using Machine Learning 2018 23rd International Scientific-Professional Conference on Information Technology (IT), Draško Radovanović, Božo Krstajić, *Member, IEEE*.

[20] Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques 2017, The Sixth International Conference on Data Analytics, Elsharif Elmurngi, Abdelouahed Gherbi.

[21] A Comparative Study on Fake Review Detection Techniques Vol 5, Issue 4, April 2018, International Journal of Engineering Research in Computer Science and Engineering (IJERCSE), A.Lakshmi Holla, Dr Kavitha K.S.

[22] Detection of fake online hotel reviews, 2017, 2th International Conference for Internet Technology and Secured Transactions (ICITST), Anna V. Sandifer ; Casey Wilson ; Aspen Olmsted.

[23] Opinion mining using Ontological Spam Detection, 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS), Neelam Duhan, Divya ; Mamta Mittal

[24] Research on Product Review Analysis and Spam Review Detection, 2017 4th International Conference on Signal Processing and Integrated Networks (SPIN), Shashank Kumar Chauhan, Anupam Goel, Prafull Goel, Avishkar Chauhan and Mahendra K Gurve.

[25] Support Vector Machine Equipped with Deep Convolutional Features for Product Review Classification, 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD 2017), Shangdi Sun, Xiaodong Gu.

[26] Identifying Indicators of Fake Reviews Based on Spammer's Behavior Features, 2017 IEEE International Conference on Software Quality, Reliability and Security (Companion Volume), Pan Liu^{1,2}, Zhenning(Jimmy) Xu³, Jun Ai⁴, Fei Wang⁴.

[27] Fake Product Review Monitoring and Removal for Genuine Online Product Reviews Using Opinion Mining, Volume 7, Issue 1, January 2017, International Journal of Advanced Research in Computer Science and Software Engineering, Manleen Kaur Kohli, Shaheen Jamil Khan, Tanvi Mirashi, Suraj Gupta