

Trustworthiness of User Feedback in the Context of E-Commerce

Lakshmisampanna N¹ Prof. Kavita P. Shirsat²

¹ME Scholar ²Assistant Professor

^{1,2}Department of Computer Science

^{1,2}Vidyalankar Institute of Technology, Mumbai

Abstract— Due to the tremendous development in the field of web technology doors have opened for several e-commerce websites to market and sell products online. They also provide opportunity for the users to post opinions or reviews based on their experience in buying the products. So the trend has been set by buyers to go through reviews before buying any products and make decisions accordingly. But the main question that comes into picture is whether these reviews are believable or not. Nowadays due to review's capacity to change buyer's decision, there are people or firms whose only job is to write spam reviews either to falsely defame or to promote a product which is unethical. This paper mainly deals with a trust system through which the genuineness of the review is checked, which is based on the behavior of its reviewer towards other pre-fabricated reviews. Only those proved trustworthy can publish their reviews in the webpage so that those reviews are considered for opinion mining.

Key words: Trust Reputation System, Aspect Mining, Reviews, Pre-Fabricated Feedback, Trustworthiness

I. INTRODUCTION

Nowadays Internet offers an effective platform for e-commerce, opinion sharing. There are several blogs, sites devoted only for opinion sharing i.e. views on politics, finance, religion, culture, history, sports, education etc. So enormous information is available online and mining has huge scope here. Mining of these terabytes of data is challenging task. Making use of these data is important to understand people's need, to make marketing strategies etc.

These days, online shopping i.e. purchasing the products online by placing orders in their respective websites and doing the financial transaction and getting the product delivered to our door step is the trend which is getting popular. Prior to this we had offline traditional method where purchaser who wants to make any purchase will get advice from his/her friends, family etc and decide on which product to purchase and then making a purchase by visiting the respective shop to buy it. But now, buyer typically browses through several similar products and goes through the previous buyers reviews and compares different products and then comes to a conclusion as to go for that product or not. And then buys it online.

But there will be thousands of reviews for a product, which is very difficult for a person to go through one by one. This is where mining comes into picture. Web mining or more specifically Opinion mining is that area of data mining which deals with reviews or opinions which extract the desired information from the unstructured user review.

Main question that comes next is whether these reviews which are used for mining activities are genuine? This is an important issue with respect to reviews which are present online. Trust is an important factor in any relationship, be it social. There are firms or people dedicated

to deceive others. They are called spammers. There are different kinds in spammers. But the one who writes fake review are review spammers and the unethical activity is called review spamming.

These spammers fill the websites with fake reviews i.e. writing more number of positive reviews for a product which may not be good at all, so that the buyers are encouraged to buy that product. On the other hand posting negative comments to some other rival organization's product there by building negative image in the market which in turn discourages the buyer's decision of buying it.

Fake reviews are very difficult to deal with. Because reviewers are expert in writing fake reviews so that one cannot say a particular review is a fake just by manually reading them.

The idea of this paper is to combine Opinion mining activity along with the concept called Trust. We are using Trust Reputation System (TRS) [2] and Opinion miner. TRS which will ensure trustworthy feedback is passed to opinion miner. And an Opinion miner will extract features from the review and classifies it based on the opinion obtained from the user. Finally summarization of the mining activity is done.

II. LITERATURE SURVEY

Jeyapriya and Selvi [4] created a system which does extraction of features at aspect level. Frequent Item set Mining is used to find all frequent sets using minimum support count. Sentence and Aspect orientation are implemented using Naïve Bayesian Algorithm using supervised term based approach. Sharma, Nigam and Jain [5] also dealt with opinion mining at aspect level which used dictionary based approach of the unsupervised technique for aspect orientation. WordNet is used as a Dictionary to determine the opinion words and their synonyms and antonyms. Seed list preparation is done here for storing the extracted opinions and their polarity. The polarity of the sentence is decided on the basis of majority of opinion words. Khushboo, Vekariya and Mishra [6] mainly created a mining system for sentence level. They used Supervised Term weighted approach of Naïve Bayesian Algorithm for creating the table of opinions and knowing their polarity. With respect to spam reviews, Manali S. Patil and A. M. Bagade [8] proposed a system to detect untruthful spam reviews using n-gram language model and detect review on brand spam detection using feature selection technique. Many authors such as the authors of [9],[10],[11] propose in their work several TRS architectures with different algorithms to calculate reputation score related to the product.

III. BASICS OF OPINION MINING AND TRUST

Opinion mining also called as sentiment analysis. It is a process of finding users opinion about particular topic or a

product or a problem. A topic can be anything a news article, product, movie etc.

The goal of Opinion Mining is to make computer able to understand and express emotions. A thought or attitude based on emotions instead of a reason is called sentiment.

Opinion mining is extracting people's opinion from the web. It analyzes people's opinions, appraisals, attitudes, and emotions toward organizations, entities, persons, issues, actions, topics, and their attributes [12].

Opinion is quintuple($e_j, a_{jk}, so_{ijkl}, h_i, t_l$), where e_j is the target entity, h_i is opinion holder, a_{jk} is aspect of entity t_l is the time when the opinion is expressed, so_{ijkl} is sentiment orientation of opinion holder h_i on feature a_{jk} of entity e_j at time t_l [4].

There are three different types in Opinion mining based on the level of data extraction. Document level, Sentence level and phrase level Opinion mining. Document level classifies whole document as positive document or a negative document. Sentence level opinion mining breaks the document into individual sentences and then classifies each sentence into positive or negative. Last and the most effective mining is the Phrase or Aspect level mining where each individual sentence is broken so that each aspects or features are extracted along with their respective polarity.

Trust is an important factor in any kind of relationship. It can be personal, professional, social etc. As e-commerce transactions usually involve finance, proper on time delivery of goods, maintaining the customer relationships, this Trust factor plays a vital role. People should have Trust in whichever the product, site they are dealing with. If trust is built then the reputation builds which in turn benefits the product or the website.

To define Trust it is the firm belief in the competence of an entity to act dependably, reliably and securely within a specific context [7]. TRS helps people detect the trustworthy parties and influence the buyer's decision in doing purchase.

IV. EXISTING SYSTEM

Nowadays buyer's buying decisions are dependent on reviews or feedbacks given by other customers. So there is a lot of spamming that is occurring in this area. Spammers are paid in huge amounts by some organization just to mislead the buyers. The opinion spammer's objective is to promote or to demote the reputation of some targeted products. There is a huge demand in identifying these fake reviews and reviewers. On the other hand there is lot of scope in opinion mining field. Usually huge reviews are taken as it is from the websites and mining activity is performed on them. In this collection of reviews there may be lot of spam reviews in it which may hamper the actual result. It may give excellent results but their genuineness is not guaranteed. If these online reviews on which our purchases, decisions, money depends are not real then we will definitely feel cheated.

V. PROPOSED SYSTEM

This project aims to combine concepts of trust with opinion mining. Incorporating a certain degree of trust in opinion mining may improve the quality of the reviews published in

the websites and mining those reviews will lead into getting actual opinion of the customer. We try to use TRS algorithm [2] to find genuineness of the buyer and then trustworthy user's feedback is made to publish in the respective sites. For mining we take the reviews which are obtained from TRS and true opinion for the products based on their features are found.

VI. METHODOLOGY

This section mainly dedicates itself to the whole strategy of this project.

The whole strategy can be divided into 2 modules:

- TRS module
- The Opinion miner

A. TRS Module

1) Architecture of TRS

- Initially user logs in and gives feedback and appreciation (rating).
- Next step involves concordance test which checks whether there is a match between user's textual feedback and user rating.
- Once concordance test passed, user is taken to a page containing random pre-fabricated feedbacks where user is asked to click either like or unlike button based on their opinion on those feedbacks.
- The trustworthiness of those pre-fabricated feedbacks is already calculated and stored using trust reputation algorithm. Using this user's behaviour toward other feedbacks is checked and user's trustworthiness is calculated.
- Once the user crosses the limit of untrustworthiness he is proved trustworthy. Hence we can trust that feedback.
- Storing all reviews in the database whose users are proved trustworthy.

These feedbacks are used for Opinion mining. Hence to a certain extent may not be completely, we can filter out the untrustworthy feedbacks.

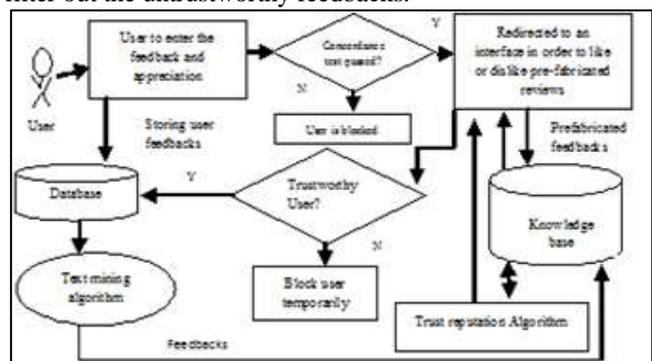


Fig. 1: Architecture of TRS

B. The Opinion Miner

Steps involved are:

- Select the product whose reviews are to be mined.
- Project aims two types of mining. User can go for an individual review mining. Where individual reviews are mined. Or we can mine all reviews together.
- Initially as part of pre-processing Parts Of Speech (POS) tagger applied which tags all the words in a sentence.

- We extract the feature, Opinion based on POS tagger where Noun is for feature and Adjective stands for opinion.
- Opinions are checked against SentiWordNet, a lexical resource for determining the polarity of these opinions.
- Finally result is displayed.

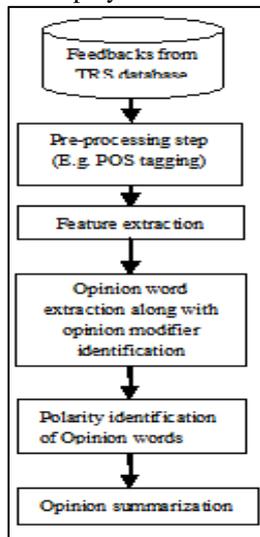


Fig. 2: Architecture of the Opinion miner

Opinion mining here in this project is considered in two types.

- Individual Review Mining
- Overall Opinion Mining

1) Individual Review Mining

In individual review mining technique each user feedback can be mined individually and result is displayed. This is a simplifying technique for the user who doesn't want to go through every single line in a feedback but requires information about what aspect is good and which aspect is not good etc.

Consider an example "this is very good phone but camera is bad" where the manual way we can identify that camera aspect has received a negative opinion where "bad" is a negative word and is determined by SentiWordNet. Whereas the aspect camera received a positive opinion "Good" according to the lexical resource which classifies opinion words (adjectives) into either positive or negative, I.e. SentiWordNet. Figure below shows the result of the individual mining for the same review as discussed.

Sentiment Analysis of Google Nexus 5		
Feature	Positive Polarity	Negative Polarity
camera	0	1
phone	1	0

Fig. 3: Result of individual review mining

2) Overall Opinion Mining

In this mining technique we combine all individual review mining results and we give user an overall result combining all reviews for the particular product. This is for those users who cannot go through individual reviews one by one. All reviews will be compiled together and result is displayed.

Figure below is showing the result we got once we selected overall mining option for our product Google nexus 5.

Sentiment Analysis of Google Nexus 5		
Feature	Positive Polarity	Negative Polarity
battery	2	3
camera	1	4
HD	1	0
phone	6	0
processor	0	1
product	1	0
resolution	1	0
screen	3	0
tethering	1	0
WiFi	1	0

Fig. 4: Result of Overall opinion mining for a single product

VII. UNTRUSTWORTHY USER HANDLING MECHANISM

Any user whose feedback is proved untrustworthy has a different handling mechanism. Initially the user gets blocked temporarily (say here 10days), but his id will enter blocked list database permanently. Next whenever the user tries to login his blocked list is checked. The temporary blocking status changes when the blocked day crosses 10 days. And system allows blocked list user to login. If blocked user tries to access within that temporary blocked number of days then access will be denied.

If user tries to login after say 10days i.e. once he clears the temporary blocked number of days then he/she has to be allowed to login. Now this project has a different way to handle them. As they login their entry will be cross checked with respect to the blocked user database. For blocked user there will be an entry in the database. So once that is confirmed then as usual the feedback undergoes concordance check. If it fails there again it is blocked.

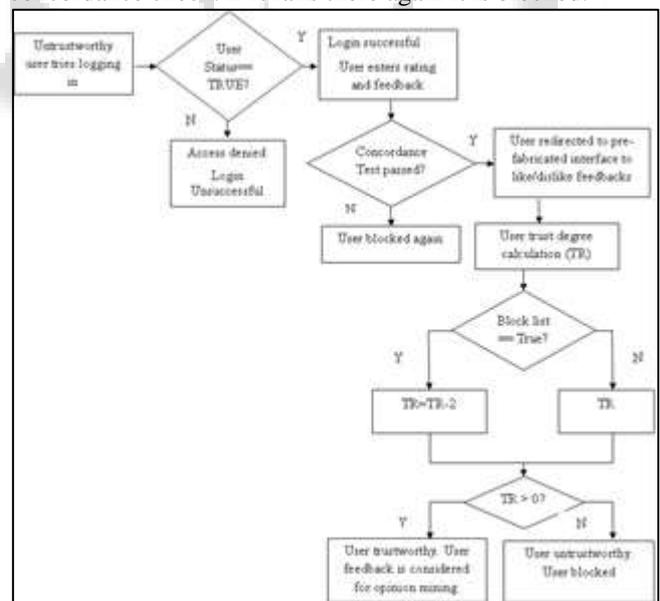


Fig. 5: Flowchart for untrustworthy user handling mechanism

Else it goes to the pre-fabricated interface where the prefabricated feedbacks have to be validated. This process requires trust algorithm to find the user trust degree. Once user trust degree is found we simply reduce a constant value from the user trust degree (here say 2). Now again the new trust degree is checked and if the degree is still positive then we publish them considering them trustworthy by updating the trust value of the user. If the new calculated

trust degree is a negative value again the user is proved untrustworthy and blocked again.

If user is trustworthy then as usual the user degree is updated, user's feedback is published in the website and also user's feedback is considered for opinion mining technique. If user is proved untrustworthy, once again he will be blocked. User degree is updated, user feedback is stored in the database for knowledge base but it won't be published in the website and hence won't be considered for mining activity.

VIII. COMPARISON BETWEEN TWO PRODUCTS

This is a feature that is included in the project which helps the user to get a better idea about the product that he is going to buy. Consider if buyer has shortlisted on what product he/she is interested in and is unable to decide on which among two to decide. Comparison feature can come to the rescue. Here it will not compare features of two products like what feature one product has one which it doesn't have. Here it takes the help of Opinion mining results that it has calculated.

Initially the buyer will select two products for comparison and upon selecting the products the system internally loads all its opinion mining results with respect to its reviews. The comparison takes place with respect to the positive reviews against the all common feature between two product that is been considered. It is plotted as a graph.

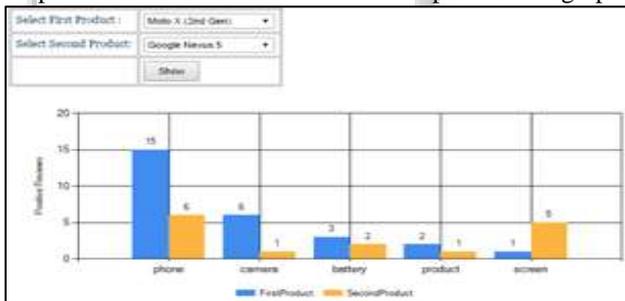


Fig. 6: Comparison graph between two products

Above figure shows the comparison between products Moto X and Google nexus 5. Only positive reviews are considered and only common aspects between both the products are considered. This helps user better to make decisions between the products.

IX. RESULT ANALYSIS

For result analysis we have taken mobile phones as review objects and considered reviews of it as inputs for calculating the parameters of Opinion mining

The performance of the system is evaluated. Precision, Recall and F-measure are the parameters used in the opinion mining system for evaluation.

Precision is the measure of retrieved instances that are relevant. Recall is the fraction of relevant instances that are retrieved. F-measure is a measure of test's accuracy.

Precision, Recall and F-measure are defined as follows as in [4],

$$\text{Precision} = \frac{|\text{Extracted values} \cap \text{True values}|}{|\text{Extracted values}|}$$

$$\text{Recall} = \frac{|\text{Extracted values} \cap \text{True values}|}{|\text{True values}|}$$

$$\text{F-measure} = \frac{2 * \text{Recall} * \text{Precision}}{(\text{Recall} + \text{Precision})}$$

To calculate these measures, true values in reviews are identified manually. Then system mines the aspects and opinion (extracted values). At this stage we get 89.8% of Accuracy and Precision 92%, Recall 94.8%, F-measure 93.2%.

Trust is not a quantitative measure and hence we cannot measure it. But even if the result of miner is not 100% accurate, we can say that the reviews considered for mining were to a certain degree is genuine, which is the main aim of this paper.

X. CONCLUSION AND FUTURE WORK

Opinion mining with the trust reputation can be a great boon to the customers who usually depend on doing e-commerce applications. But opinion mining is not an easy task to perform as it deals with natural Language Processing. To understand what people write and understand, the emotion or the sentiment behind it is a challenging task. Here this paper deals with simple opinion mining procedure by giving priority to genuineness of the feedbacks.

In future work, we plan to analyze the result more critically by posting many more reviews of various kinds and then check how the TRS and Opinion miner work. This paper has limited itself with respect to the number of reviews for result analysis. But we believe system should work more effectively with more and more reviews.

In opinion mining we have not used any popular effective algorithm. So future scope can be, improving the opinion miner using some effective algorithms. So that it becomes more efficient in handling feedbacks.

REFERENCES

- [1] Mita K.Dalal and Mukesh A.Zaveri, "Semisupervised learning based opinion Summarization and classification for Online Product reviews", Research Article, Hindawi publications corporation, Volume 2013.
- [2] Hasnae RAHIMI and Hanan EL BAKKALI, "A New Reputation Algorithm for Evaluating Trustworthiness in E-Commerce Context", IEEE 2013.
- [3] G. Angulakshmi, Dr.R.Manicka Chezian, "An Analysis on Opinion Mining: Techniques and Tools", International Journal of Advanced Research in Computer and Communication Engineering, vol. 3, Issue 7, July 2014.
- [4] A.Jeyapriya, C.S Kanimozi selvi, "Extracting Aspects and Mining Opinions in Product Reviews using Supervised Learning Algorithm", IEEE sponsored 2nd international conference on electronics and communication systems(ICECS '2015).
- [5] Richa Sharma, Shweta Nigam and Rekha Jain, "Mining of product reviews at Aspect level", International Journal in Foundations of Computer Science & Technology (IJFCST), Vol.4, No.3, May 2014.
- [6] Trivedi Khushboo N, Swati K. Vekariya, Prof.Shailendra Mishra, "Mining of Sentence Level Opinion Using Supervised Term Weighted Approach of Naïve Bayesian Algorithm", Trivedi Khushboo N et al, Int. J. Computer Technology & Applications, Vol 3(3), 987-991.

- [7] A. Josang (1996). The right type of trust for distributed systems. In C Meadows, editor, Proc. of the New Security Paradigms Workshop. ACM, 1996.
- [8] Manali S.Patil and A.M Bagade, "online Review Spam Detection using language model and feature selection", IJCA 2012.
- [9] A. Josang R. Hayward Simon Pope: Trust Network Analysis with Subjective Logic. Proceedings of the Second International Conference on Emerging Security Information, Systems and Technologies (Securware 2008), Cap Esterel, France, August 2008.
- [10] S. Steinbrecher, S. Grob, and M. Meichau: Jason: A Scalable Reputation System for the semantic Web. In the proceedings of IFIP International Federation for Information Processing 2009.
- [11] A. Josang and Ismail, Roslan and Boyd, Colin. A survey of trust and reputation systems for online service provision. Decision Support Systems 43(2):pp. 618-644. (2007).
- [12] Bing Liu (2012), 'Sentiment Analysis and Opinion Mining', Synthesis Lectures on Human Language Technologies, Morgan & Claypool Publishers.

