EFFICIENTLY ANALYZING AND DETECTING FAKE REVIEWS

Chintaram Varma[1], Vishal Das[2], Aniket Majalkar[3], Praful Jangam[4], Pravin Jangid[5], Ragini Mishra[6]

[1,2,3,4] *Student, Dept of Computer engineering, SLRTCE, Maharashtra, India*
[5,6] *Assistant Professor, Dept of Computer engineering, SLRTCE, Maharashtra, India*

Abstract *- a robust Trust Reputation Systems (TRS) provides actionable information to support the relying parties taking the right decision in any of the electronic transaction. In fact, as security providers in e-services, TRS have to faithfully calculate the most trustworthy score for a targeted product or service. Thus, TRS must rely on a robust architecture and suitable algorithms that are able to select, store, generate and classify scores and feedbacks. In this work, we propose a new architecture for TRS in e-commerce application which includes feedbacks' analysis in its treatment of scores. In fact, this architecture is based on an intelligent layer that proposes to each user (i.e. "feedback provider") who has already given his recommendation, a collection of prefabricated feedbacks summarizing other users' textual feedbacks. A proposed algorithm is used by this architecture in order to calculate the trust degree of the user and the feedback's trustworthiness and generates the global reputation score of the product.*

***Keywords*—prefabricated; concordance; trust reputation system; trust degree;**

1. INTRODUCTION

Trust is an important factor in any social aspects and especially in commerce transactions. In the e-commerce context, there is a lack of trust assessment. Though electronic signatures cryptography and certificates help users to make the transaction more secure, they are insufficient to build a trustful reputation about a particular product or a service. Hence, users are not able to conceive a reputation for the product without any additional help.

E-commerce users focus on users' opinions for a product or service in order to conceive their own trust and reputation experience. Users believe in their common interest which is to know about the trustworthiness of the transaction and product. Therefore feedbacks or reviews and any other information provided by users are very useful for the trust reputation assessment. But the reliability of this information needs to be verified.

TRS is indeed essential mechanisms that aim to detect malicious interventions of users whose intention is to falsify the Reputation score of a product positively or negatively. In the literature, there are many works such as that proposes algorithms for calculating a reputation or defining a specific set of possible reputations or ratings. However, few of them such as have been devoted to the semantic analysis of textual feedbacks in order to generate a most trustful trust degree of the user.

2. LITERATURE REVIEW

Now we will look at the literature survey of the project and what all projects exist prior and were actually used in the market which the makers of this project took the inspiration from and thus decided to proceed with the project covering with the problem statement. Many works propose TRS architecture together with different algorithms to calculate the reputation score related to a product. Nevertheless, few research works on TRS has considered the semantic analysis of feedbacks and especially the trust degree of the user in the calculus of product's trusts score. To give more complex reputation, such as the inclusion of the trust degree of the user in the calculus of a trustful reputation score for a product, the update of the trust degree of the user, the novelty of the rating and especially the feedback, the concordance between the given rating which is a scalar value and the textual review associated to it unlike those TRS, our design deals with these issues and uses a reputation algorithm that includes semantic analysis of textual feedbacks so as to calculate the trust degree of the user.

For example, the authors propose a method that uses subjective logic in order to examine trust network (TNA-SL). Hence, this method aims to model in a simple way the relationship between different agents. A single arc means a single trust relationship between two nodes A and B [A:B] meaning that A trusts B. However, this trust should have degrees that can represent how much A trusts B. This issue is not taken into account in the paper. In this paper, for each user

Who wants to leave a textual feedback (semantic review) and a rating (appreciation) we examine his attitude toward a number of short and selected prefabricated feedbacks stored by product in the knowledge base. This user's review is going to be reached by any other user. Then, we suppose that we have a path relaying all the users. Finally, we need to know the trust degree of the user and determine the trust degree of the review or feedback.

3. SENTIMENT ANALYSIS

Sentiment Analysis tells about the attitude of a speaker about a particular topic. It is also known as opinion mining and is a process of determining positivity, negativity and neutrality of a piece of writing. For this purpose text analysis, natural language processing, computational linguistics etc. are used.

4. EXISTING SYSTEM

4.1. Product Aspect Ranking and Its Application:

Zheg-Jun-Zha [1] first identify product aspects by a shallow dependency parser and determine consumer opinions on these aspects via a sentiment classifier. Then [1] develop a probabilistic aspect ranking algorithm to infer the importance of aspects by simultaneously considering aspect frequency and the influence of consumer opinions given to each aspect over their overall opinions. The experimental results on a review corpus of 21 popular products in eight domains demonstrate the effectiveness of the proposed approach. Moreover, we apply product aspect ranking [2] to two real-world applications, i.e., document-level sentiment classification and extractive review summarization, and achieve significant performance improvements, which demonstrate the capacity of product aspect ranking [2] in facilitating real-world applications.

4.2. Fake Review N Brand Spam Detection Using J48 Classifier:

In this paper Sushant Kokate [3] and Bharat Tidke [3] proposed the method to recognizing the untruthful reviews that are given by the users which having distinct semantic content based on sentiment analysis as the reviews of movies. In this paper [3] represent to detect the spam untruthful reviews of movies [4]. For this classification they used J48 classifier. Through which they have generated ARFF from the distinct features to detecting the untruthful reviews. Using Support Count in Association Rules they further detect Brands in Fake Reviews.

4.3. Fuzzy Based Sentiment Analysis of Online Product Reviews Using Machine Learning Technique:

Haseena Rehmath P [5] and Tanvir Ahmad, PhD [6].In their proposed system it can be used for binary as well as fine-grained sentiment classification of user reviews. The proposed technique utilizes fuzzy functions to emulate the effect of various linguistic hedges such as dilators, concentrator and negation on opinionated phrases that make the system more accurate in sentiment classification and summarization of users' reviews [7-10].Experimental evaluation indicates the system can perform the sentiment analysis with an accuracy of 93.85%.

4.4. Sentiment Analysis and Summarization of Twitter Data:

Seyed-Ali Bahrainian, Andreas Dengel [11].In this paper they have introduces a novel solution to target-oriented (i.e. aspect based) sentiment summarization and SA of short informal texts with a main focus on Twitter posts known as "tweets". They compare different algorithms and methods for SA polarity detection and sentiment summarization [12]. They have shown that their hybrid polarity detection system [12] not only outperforms the unigram state-of-the-art baseline, but also could be an advantage over other methods when used as a part of a sentiment summarization system. Additionally, they have illustrate that their SA and summarization system exhibits a high performance with various useful functionalities and features.

5. PROPOSED SYSTEM

At first, the user provides an appreciation (rating) then a textual feedback on a particular product. The TRS requires a text mining algorithm which aims to get the given information and verify the concordance between the user's given appreciation and the textual feedback, so as to avoid and eliminate any contradiction.

Once the concordance verified, we redirect the user to an interface of selected pre-fabricated feedbacks. So as long as we add feedbacks in the data base of origin, a text mining algorithm is going to make pre-fabricated feedbacks with different categories and fill out the knowledge base (Fig. 1 shows the architecture). The text mining algorithm would contain a part of learning in order to automatically fill out the knowledge base. The user is invited to like or dislike each feedback of the dis-played selection. Each review or feedback has already a score of trustworthiness which shows or represents the trust degree of the user who is the provider of the feedback. The user can select the number of short feedbacks like and dislike (min=4 and max=10).

Now the proposed reputation algorithm gets the user's view on each review (like/dislike) in addition to the trustworthiness degree of the liked/disliked feedback and uses them to generate a trust degree for the user. The architecture hereafter represents the connection between the e-commerce application and the solicited TRS showing the intervention of both the text mining and the Reputation algorithm.

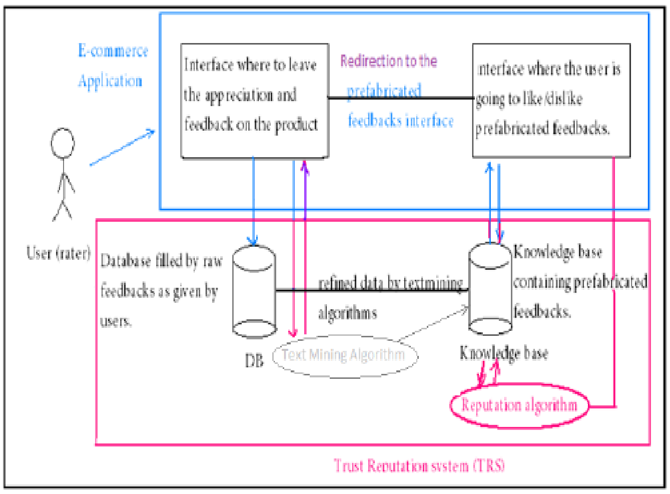

Fig. 1. Trust Reputation System Architecture

6. PROPOSED REPUTATION ALGORITHM

Before giving details on the approach of the Reputation algorithm, we will start first with giving an overview on the steps of the algorithm:

1. Verify the concordance between the appreciation and the textual feedback.

2. Display to the user a selection of the most recent pre-fabricated feedbacks with different types (freshness of feedbacks), if the concordance is verified. This selection of feedbacks is to be liked or dis-liked by the user.

3. Fetch data out of the database regarding the trustworthiness of the liked or disliked feedback and the trust degree of the user.

4. Generate / update the trust degree of the user using the trustworthiness of the feedback and the user's choice (like/dislike).

5. Standardize the trust degree of the user in order to respect the threshold [-10, 10].

6. Generate the global trust score of the product using the user's trust degree as a coefficient.

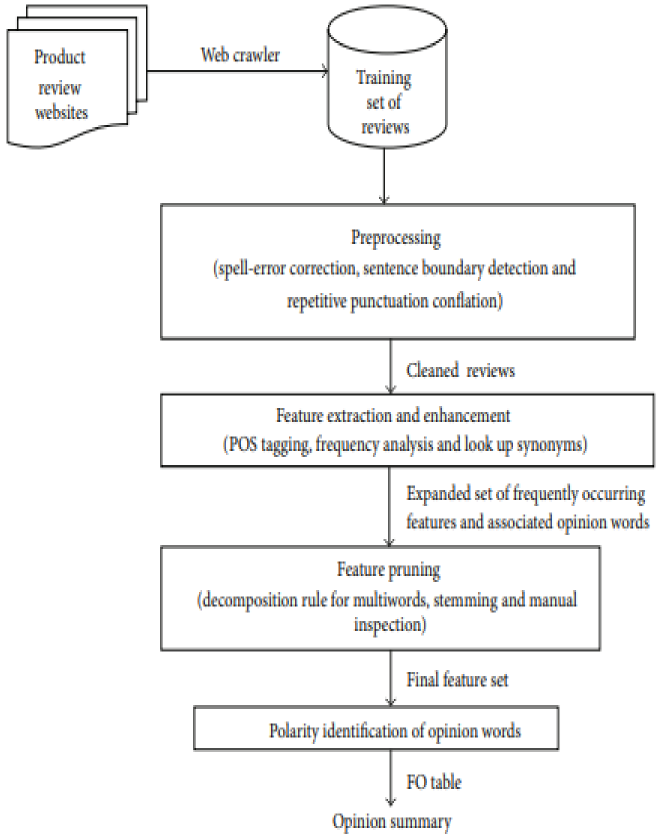

Fig 2: Trust Reputation System Flow diagram

7. SCREENSHOTS

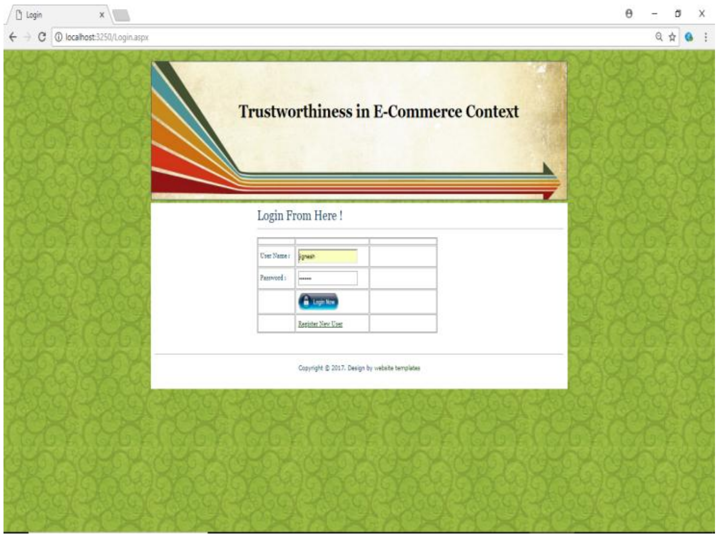

Fig 3: Login Page

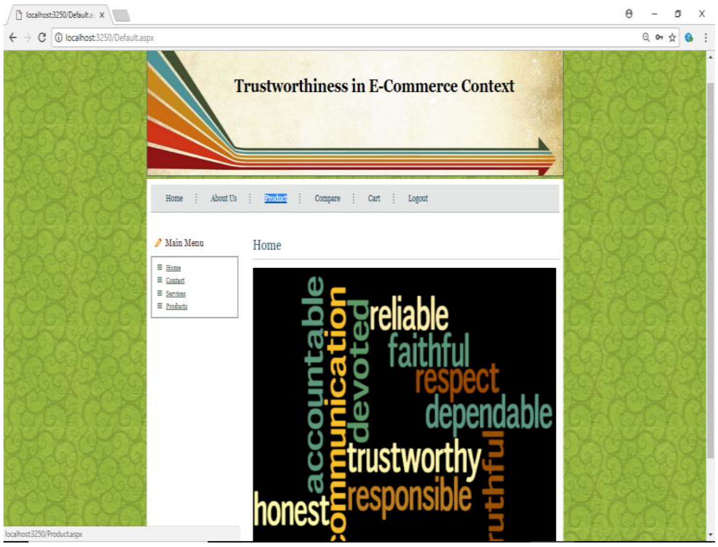

Fig 4: Home Page

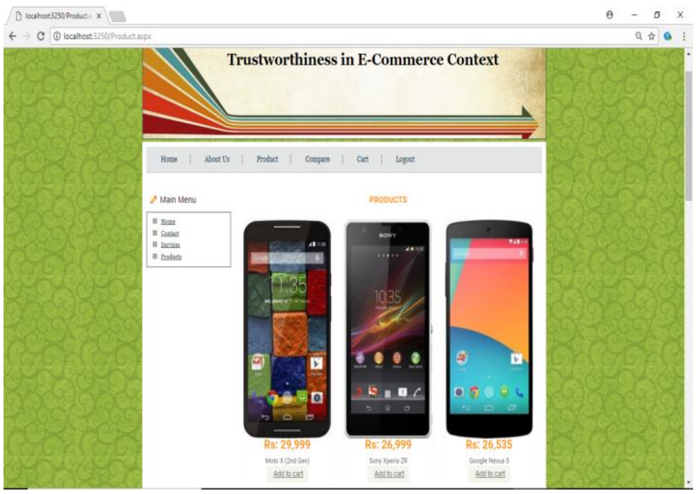

Fig 5: Product Page

Selecting a particular product to give feedback on:

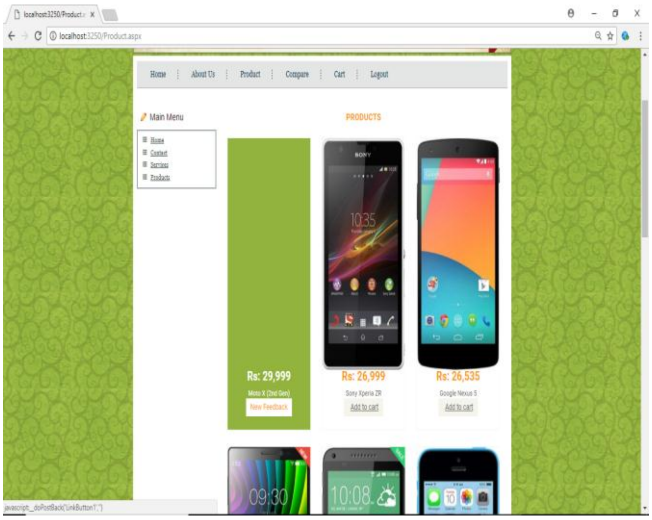

Fig 6: Product Page

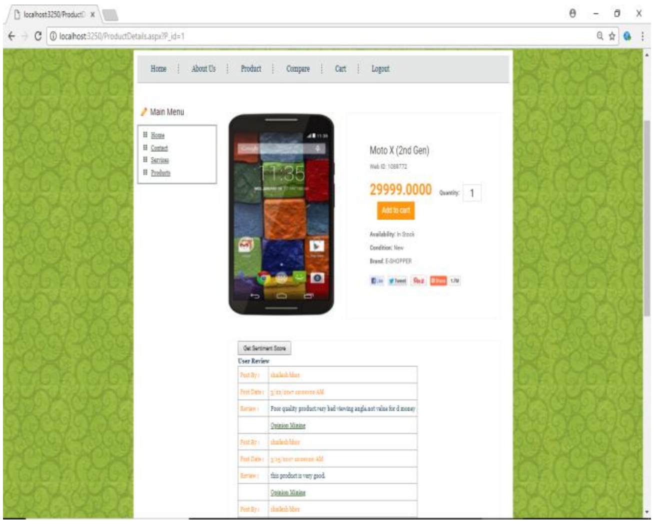

Fig 7: Particular Product Page

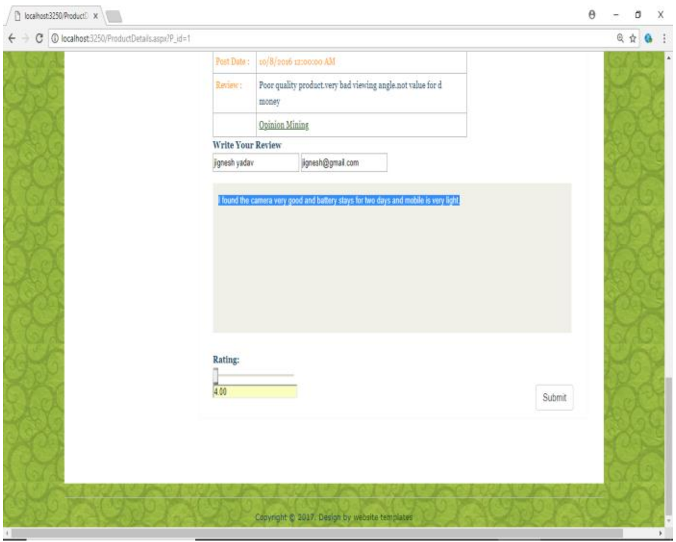

Fig 8: Review Section

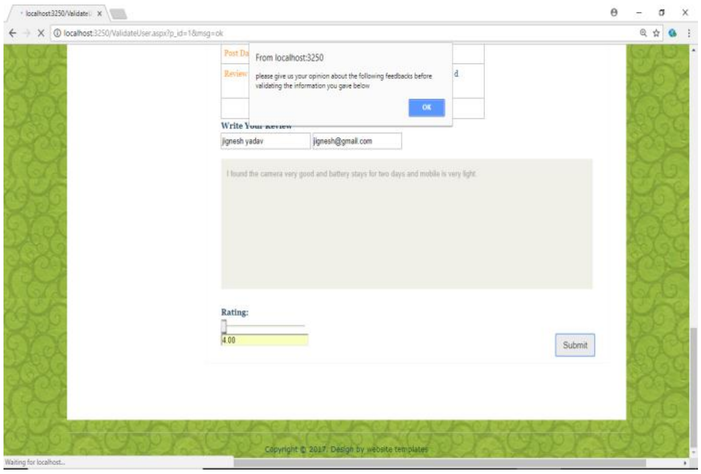

Fig 9: Concordance Verify

In the above diagram if rating and textual feedback were not in concordance then user would have been blocked for some days. Since Concordance was verified system is redirecting reviewer to another page where reviewer has to like or dislike the prefabricated feedback.

Fig 10: Prefabricated Feedback 1

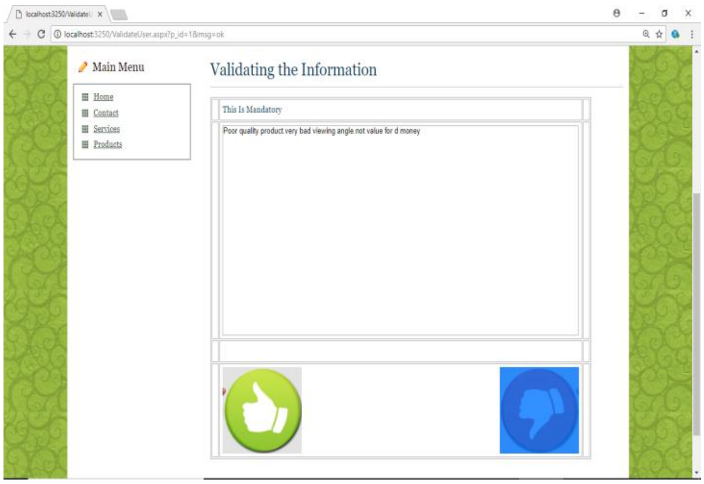

Fig 11: Prefabricated Feedback 2

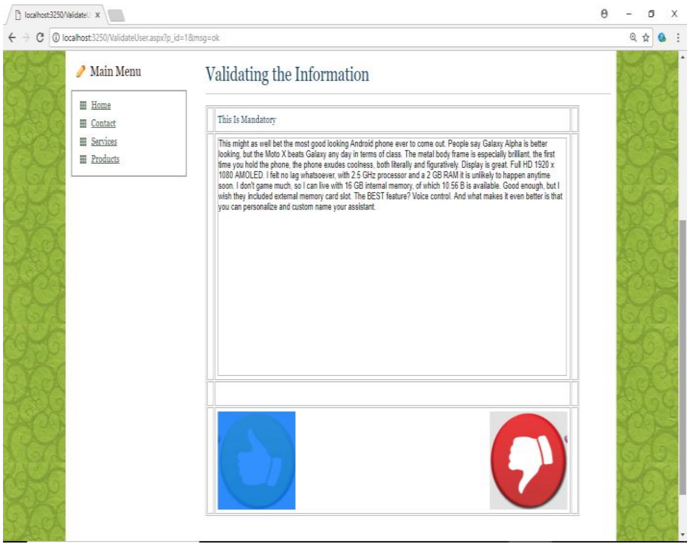

Fig 12: Prefabricated Feedback 3

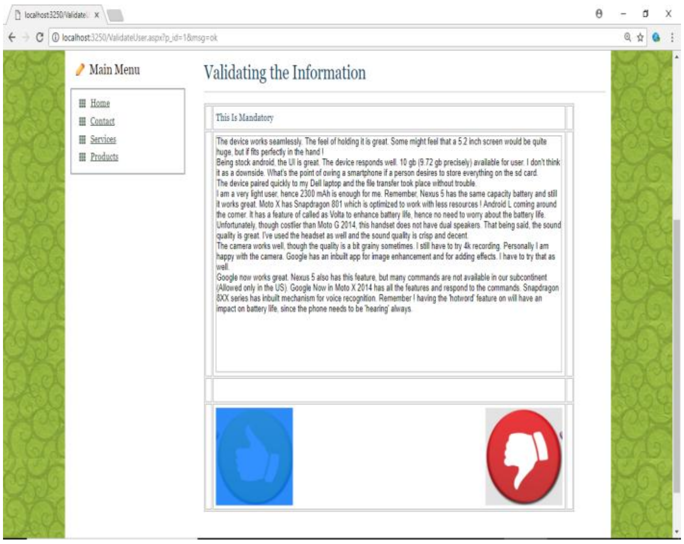

Fig 13: Prefabricated Feedback 4

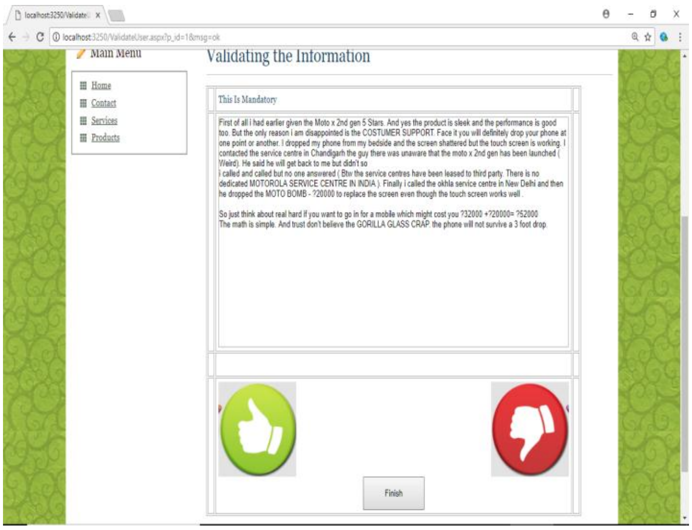

Fig 14: Prefabricated Feedback 5

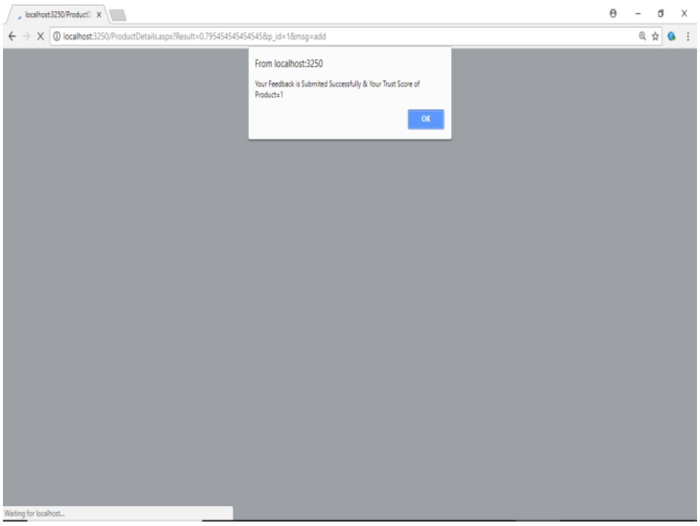

Fig 15: Feedback Acceptance

Here, feedback given by user has been accepted as shown in the above image but if it was not genuine then user would have been blocked temporarily.

8. CONCLUSION

This system detects fake reviews and blocks the user temporarily. It helps the online retailer to make a business decision. Also, online buyers find genuine reviews on websites and they can decide which product to buy. Thus this system makes e-commerce trustworthy.

FUTURE SCOPE

We can also use the live data from different websites for the same. Global trust factor calculated here can be used further for different purpose.

REFERENCES

[1] Zheng-Jun Zha, Member, IEEE, Jianxing Yu, Jinhui Tang, Member, IEEE, Meng Wang, Member, IEEE, and Tat-Seng Chua

[2] H. Wang, Y. Lu, and C. X. Zhai, “Latent aspect rating analysis on review text data: A rating regression approach,” in Proc. 16th ACM SIGKDD, San Diego, CA, USA, 2010, pp. 168–176. [35] T. Wilson, J. Wiebe, and P. Hoffmann, “Recognizing cont

[3] Sushant Kokate and Bharat Tidke (Department of Computer Engineering, Flora Institute of Technology, Pune, India)

[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] Haseena Rehmath P (M.Tech Scholar, Dept. of Computer Science and Engineering, Al-Falah School of Engineering Dauj, Haryana, India)

[6] Tanvir Ahmad, PhD(H.O.D. Dept. Of Computer Engineering Jamia Millia Islamia, New Delhi, India)

[7] A. Kennedy and D. Inkpen, "Sentiment classification of movie reviews using contextual valence shifters", Computational Intelligence, vol. 22, no. 2, pp. 110–125, 2006.

[8] L. Polanyi and A. Zaenen, "Contextual valence shifters", in Computing Attitude and Affect in Text: Theory and Applications, vol. 20 of The Information Retrieval Series, pp. 1–10, 2006.

[9] A.D. Vo and C.Y. Ock, "Sentiment classification: a combination of PMI, SentiWordNet and fuzzy function", in Proceedings of the 4th International Conference on Computational Collective Intelligence Technologies and Applications (ICCCI '12), vol. 7654, part 2 of Lecture Notes in Computer Science, pp. 373–382, 2012.

[10] S. Nadali, M. A. A. Murad, and R. A. Kadir, "Sentiment classification of customer reviews based on fuzzy logic", in Proceedings of the International Symposium on Information Technology (ITSim'10), pp. 1037–1044, mys, June 2010.

[11] Seyed-Ali Bahrainian, Andreas Dengel (Computer Science Dept., University Of Kaiserslautern, Germany Knowledge Management Dept., DFKI, Kaiserslautern, Germany {seyed_ali.bahrainian, andreas.dengel}@dfki.uni-kl.de).

[12] Bahrainian, S.A., Dengel, A., Sentiment Analysis using Sentiment Features, In the proceedings of WPRSM Workshop and the Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence, Atlanta, USA, 2013.